

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

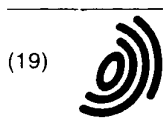
Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**



Europäisches Patentamt
European Patent Office
Office européen des brevets



(11) **EP 1 130 094 A2**

(12) **EUROPEAN PATENT APPLICATION**

(43) Date of publication:
05.09.2001 Bulletin 2001/36

(21) Application number: **00114089.6**

(22) Date of filing: **07.07.2000**

(51) Int Cl.7: **C12N 15/12, C12N 15/11,
C12N 15/10, C12N 15/70,
C12N 15/85, C12N 5/10,
C12N 1/21, C07K 14/47,
C07K 16/18, C12Q 1/68**

(84) Designated Contracting States:
**AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU
MC NL PT SE**
Designated Extension States:
AL LT LV MK RO SI

(30) Priority: **08.07.1999 JP 19448699
11.01.2000 JP 2000118774
02.05.2000 JP 2000183765**

(71) Applicant: **Helix Research Institute
Kisarazu-shi, Chiba 292-0812 (JP)**

(72) Inventors:
• **Ota, Toshio**
Fujisawa-shi, Kanagawa 251-0042 (JP)
• **Nishikawa, Tetsuo**
Tokyo 173-0013 (JP)
• **Isogai, Takao**
Inashiki-gun, Ibaraki 300-0303 (JP)
• **Hayashi, Koji**
Ichihara-shi, Chiba 299-0125 (JP)
• **Ishii, Shizuko**
Kisarazu-shi, Chiba 292-0812 (JP)

• **Kawai, Yuri**
Kisarazu-shi, Chiba 292-0812 (JP)
• **Wakamatsu, Ai**
Kisarazu-shi, Chiba 292-0014 (JP)
• **Sugiyama, Tomoyasu**
Kisarazu-shi, Chiba 292-0045 (JP)
• **Nagai, Keiichi**
Higashiyamato-shi, Tokyo 207-0022 (JP)
• **Kojima, Shinichi**
Kisarazu-shi, Chiba 292-0052 (JP)
• **Otsuki, Tetsuji**
Kisarazu-shi, Chiba 292-0055 (JP)
• **Koga, Hisashi**
Kisarazu-shi, Chiba 292-0055 (JP)

(74) Representative: **VOSSIUS & PARTNER
Siebertstrasse 4
81675 München (DE)**

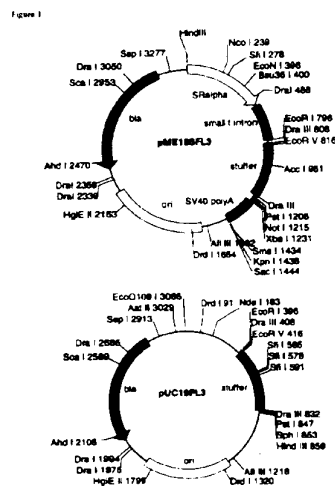
Remarks:

The sequence listing, which is published as annex to the application documents, was filed after the date of filing. The applicant has declared that it does not include matter which goes beyond the content of the application as filed.

(54) **Primers for synthesizing full length cDNA clones and their use**

(57) Primers for synthesizing full length cDNAs and their use are provided.

830 cDNA encoding a human protein has been isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA have been determined. Furthermore, primers for synthesizing the full length cDNA have been provided to clarify the function of the protein encoded by the cDNA. The full length cDNA of the present invention containing the translation start site provides information useful for analyzing the functions of the protein.



DescriptionFIELD OF THE INVENTION

5 **[0001]** The present invention relates to a polynucleotide encoding a novel protein, a protein encoded by the polynucleotide, and new uses of these.

BACKGROUND OF THE INVENTION

10 **[0002]** Currently, the sequencing projects, the determination and analysis of the genomic DNA of various living organisms have been in progress all over the world. The whole genomic sequences of more than 10 species of prokaryotes, a lower eukaryote, yeast, and a multicellular eukaryote, *C. elegans* are already determined. As to human genome, which is supposed to be composed of three thousand million base pairs, the world wide cooperative projects have been under way to analyze it, and the whole structure is predicted to be determined by the years 2002-2003. The aim of the determination of genomic sequence is to reveal the functions of all genes and their regulation and to understand living organisms as a network of interactions between genes, proteins, cells or individuals through deducing the information in a genome, which is a blueprint of the highly complicated living organisms. To understand living organisms by utilizing the genomic information from various species is not only important as an academic subject, but also socially significant from the viewpoint of industrial application.

20 **[0003]** However, determination of genomic sequences itself cannot identify the functions of all genes. For example, as for yeast, only the function of approximately half of the 6000 genes, which is predicted based on the genomic sequence, was able to be deduced. As for human, the number of the genes is predicted to be approximately one hundred thousand. Therefore, it is desirable to establish "a high throughput analysis system of the gene functions" which allows us to identify rapidly and efficiently the functions of vast amounts of the genes obtained by the genomic sequencing.

25 **[0004]** Many genes in the eukaryotic genome are split by introns into multiple exons. Thus, it is difficult to predict correctly the structure of encoded protein solely based on genomic information. In contrast, cDNA, which is produced from mRNA that lacks introns, encodes a protein as a single continuous amino acid sequence and allows us to identify the primary structure of the protein easily. In human cDNA research, to date, more than one million ESTs (Expression Sequence Tags) are publicly available, and the ESTs presumably cover not less than 80% of all human genes.

30 **[0005]** The information of ESTs is utilized for analyzing the structure of human genome, or for predicting the exon-regions of genomic sequences or their expression profile. However, many human ESTs have been derived from proximal regions to the 3'-end of cDNA, and information around the 5'-end of mRNA is extremely little. Among these human cDNAs, the number of the corresponding mRNAs whose encoding protein sequences are deduced is approximately 7000, and further, the number of full-length therein is only 5500. Thus, even including cDNA registered as EST, the percentage of human cDNA obtained so far is estimated to be 10-15% of all the genes.

35 **[0006]** It is possible to identify the transcription start site of mRNA on the genomic sequence based on the 5'-end sequence of a full-length cDNA, and to analyze factors involved in the stability of mRNA that is contained in the cDNA, or in its regulation of expression at the translation stage. Also, since a full-length cDNA contains ATG, the translation start site, in the 5'-region, it can be translated into a protein in a correct frame. Therefore, it is possible to produce a large amount of the protein encoded by the cDNA or to analyze biological activity of the expressed protein by utilizing an appropriate expression system. Thus, analysis of a full-length cDNA provides valuable information which complements the information from genome sequencing. Also, full-length cDNA clones that can be expressed are extremely valuable in empirical analysis of gene function and in industrial application.

40 **[0007]** In particular, human secretory proteins or membrane proteins are would be useful by itself as a medicine like tissue plasminogen activator (TPA), or as a target of medicines like membrane receptors. In addition, genes for signal transduction-associated proteins (protein kinases, etc.), glycoprotein-associated proteins, transcription-associated proteins, and disease-associated proteins form a gene group rich in genes whose relationships to human diseases have been elucidated.

45 **[0008]** Therefore, it has great significance to isolate novel full-length cDNA clones of human, only few of which has been isolated. Especially, isolation of a novel cDNA clone encoding a secretory protein or membrane protein is desired since the protein itself would be useful as a medicine, and also the clones potentially include a gene associated with diseases. In addition, genes encoding proteins that are associated with signal transduction, glycoprotein, transcription, or diseases are expected to be useful as target molecules for therapy, or as medicines themselves. These genes form a gene group predicted to be strongly associated with diseases. Thus, identification of the full-length cDNA clones encoding those proteins has great significance.

SUMMARY OF THE INVENTION

[0009] An objective of the present invention is to provide a primer that enables synthesizing polynucleotide from human, the resulting polynucleotide or its clone, and a protein encoded by the polynucleotide.

[0010] The inventors have developed a method for efficiently cloning a human full-length cDNA that is predicted by the ATGpr etc. to be a full-length cDNA clone, from a full-length-enriched cDNA library that is synthesized by the oligo-capping method. Then, the inventors determined the nucleotide sequence of the obtained cDNA clones from both 5'- and 3'- ends. By utilizing the sequences, the inventors selected clones that were expected to contain a signal by the PSORT (Nakai K. and Kanehisa M. (1992) Genomics 14: 897-911), and obtained clones that contain a cDNA encoding a secretory protein or membrane protein. Moreover, the inventors specifically selected full-length cDNA clones that encode secretory or membrane proteins, signal transduction-associated proteins, glycoprotein-associated proteins, transcription-associated proteins, or disease-associated proteins from clones homologous to the clones in the Swiss-Prot (http://www.ebi.ac.uk/ebi_docs/SwissProt_db/swishhome.html) according to the keywords of SwissProt.

[0011] The full-length cDNA clones of the present invention have high fullness ratio since these were obtained by the combination of (1) construction of a full-length-enriched cDNA library that is synthesized by the oligo-capping method, and (2) a system in which fullness ratio is evaluated from the nucleotide sequence of the 5'-end (in this system, clones are selected based on the estimation by the ATGpr, following the removal of sequences judged not to be full-length when compared with ESTs). However, the primers of the present invention enable obtaining full-length cDNA easily without any special methods mentioned above.

[0012] Homology analysis in which the analysis is carried out against a non-full-length cDNA fragment to postulate the function of a protein encoded by said fragment, is being commonly performed. However, since such analysis is based on the information of the fragment, it is not clear as to whether this fragment corresponds to a part that is functionally important in the protein. In other words, the reliability of the homology analysis based on the information of a fragment is doubtful, as information relating to the structure of the whole protein is not available. However, the homology analysis of the present invention is conducted based on the information of a full-length cDNA comprising the whole coding region of the cDNA, and therefore, the homology of various portions of the protein can be analyzed. Hence, the reliability of the homology analysis has been dramatically improved in the present invention.

[0013] The inventors completed the invention by finding that it is possible to synthesize a novel full-length cDNA by using the combination of a primer that is designed based on the nucleotide sequence of the 5'-ends of the selected full-length cDNA clones and any of an oligo-dT primer or a 3'-primer that is designed based on the nucleotide sequence of the 3'-ends of the selected clones.

[0014] Thus, the present invention relates to primers described below, a method for synthesizing a polynucleotide using the primers, and polynucleotides obtained by the method.

[0015] First, the present invention relates to

(1) use of an oligonucleotide as a primer for synthesizing the polynucleotide comprising the nucleotide sequence set forth in any one of SEQ ID NOs: 1-829 and 2545, or the complementary strand thereof, wherein said oligonucleotide is complementary to said polynucleotide or the complementary strand thereof and comprises at least 15 nucleotides;

(2) a primer set for synthesizing polynucleotides, the primer set comprising an oligo-dT primer and an oligonucleotide complementary to the complementary strand of the polynucleotide comprising the nucleotide sequence set forth in any one of SEQ ID NOs: 1-829 and 2545, wherein said oligonucleotide comprises at least 15 nucleotides; and

(3) A primer set for synthesizing polynucleotides, the primer set comprising a combination of an oligonucleotide comprising a nucleotide sequence complementary to the complementary strand of the polynucleotide comprising a 5'-end nucleotide sequence and an oligonucleotide comprising a nucleotide sequence complementary to the polynucleotide comprising a 3'-end nucleotide sequence, wherein said oligonucleotides comprise at least 15 nucleotides and wherein said combination of 5'-end nucleotide sequence / 3'-end nucleotide sequence is selected from the combinations of 5'-end nucleotide sequence / 3'-end nucleotide sequence set forth in the SEQ ID NOs in Table 1.

[0016] Table 1 shows names of clones obtained in the examples described later, comprising the polynucleotide of the present invention (830 clones), names of nucleotide sequences at the 5'-end and 3'-end of the full-length cDNA, and their corresponding SEQ ID NOs. A blank indicates that the of the 3'-end sequence corresponding to the 5'-end sequence has not been determined the same clone.

[0017] The SEQ ID NO of a 5'-end sequence is shown on the right side of the name of the 5'-end sequence, and the SEQ ID NO of a 3'-end sequence is shown on the right side of the name of the 3'-end sequence.

EP 1 130 094 A2

Table 1

Correspondence between names of clone and the sequence name, and the SEQ ID.					
	Name of clone	Name of 5'-sequence	SEQ ID	3 Name of 3'-sequence	SEQ ID
5	BNGH41000020	F-BNGH41000020	1		
	BNGH41000087	F-BNGH41000087	2		
	BNGH41000091	F-BNGH41000091	3		
10	HEMBA1000006	F-HEMBA1000006	4	R-HEMBA1000006	830
	HEMBA1000121	F-HEMBA1000121	5	R-HEMBA1000121	831
	HEMBA1000128	F-HEMBA1000128	6	R-HEMBA1000128	832
	HEMBA1000275	F-HEMBA1000275	7	R-HEMBA1000275	833
	HEMBA1000300	F-HEMBA1000300	8	R-HEMBA1000300	834
15	HEMBA1000349	F-HEMBA1000349	9	R-nnnnnnnnnnnnn	835
	HEMBA1000443	F-HEMBA1000443	10		
	HEMBA1000462	F-HEMBA1000462	11	R-HEMBA1000462	836
	HEMBA1000477	F-HEMBA1000477	12	R-HEMBA1000477	837
20	HEMBA1000590	F-HEMBA1000590	13	R-HEMBA1000590	838
	HEMBA1000634	F-HEMBA1000634	14	R-HEMBA1000634	839
	HEMBA1000671	F-HEMBA1000671	15	R-HEMBA1000671	840
	HEMBA1000713	F-HEMBA1000713	16	R-HEMBA1000713	841
	HEMBA1000732	F-HEMBA1000732	17	R-HEMBA1000732	842
25	HEMBA1000745	F-HEMBA1000745	18	R-nnnnnnnnnnnnn	843
	HEMBA1000835	F-HEMBA1000835	19		
	HEMBA1000875	F-HEMBA1000875	20	R-HEMBA1000875	844
	HEMBA1000907	F-HEMBA1000907	21		
30	HEMBA1000940	F-HEMBA1000940	22	R-HEMBA1000940	845
	HEMBA1000962	F-HEMBA1000962	23	R-HEMBA1000962	846
	HEMBA1001184	F-HEMBA1001184	24	R-HEMBA1001184	847
	HEMBA1001221	F-HEMBA1001221	25	R-HEMBA1001221	848
	HEMBA1001228	F-HEMBA1001228	26	R-HEMBA1001228	849
35	HEMBA1001272	F-HEMBA1001272	27	R-HEMBA1001272	850
	HEMBA1001296	F-HEMBA1001296	28	R-HEMBA1001296	851
	HEMBA1001297	F-HEMBA1001297	29	R-HEMBA1001297	852
	HEMBA1001390	F-HEMBA1001390	30	R-HEMBA1001390	853
	HEMBA1001563	F-HEMBA1001563	31	R-HEMBA1001563	854
40	HEMBA1001621	F-HEMBA1001621	32	R-HEMBA1001621	855
	HEMBA1001878	F-HEMBA1001878	33	R-HEMBA1001878	856
	HEMBA1001886	F-HEMBA1001886	34	R-HEMBA1001886	857
	HEMBA1002048	F-HEMBA1002048	35	R-HEMBA1002048	858
45	HEMBA1002131	F-HEMBA1002131	36	R-HEMBA1002131	859
	HEMBA1002163	F-HEMBA1002163	37	R-HEMBA1002163	860
	HEMBA1002164	F-HEMBA1002164	38		
	HEMBA1002167	F-HEMBA1002167	39	R-HEMBA1002167	861
	HEMBA1002178	F-HEMBA1002178	40	R-HEMBA1002178	862
50	HEMBA1002195	F-HEMBA1002195	41	R-HEMBA1002195	863
	HEMBA1002227	F-HEMBA1002227	42	R-HEMBA1002227	864
	HEMBA1002239	F-HEMBA1002239	43		
	HEMBA1002316	F-HEMBA1002316	44	R-HEMBA1002316	865
55	HEMBA1002420	F-HEMBA1002420	45	R-HEMBA1002420	866
	HEMBA1002421	F-HEMBA1002421	46	R-HEMBA1002421	867
	HEMBA1002524	F-HEMBA1002524	47	R-HEMBA1002524	868

EP 1 130 094 A2

Table 1 (continued)

Correspondence between names of clone and the sequence name, and the SEQ ID.					
	Name of clone	Name of 5'-sequence	SEQ ID	3 Name of 3'-sequence	SEQ ID
5	HEMBA1002551	F-HEMBA1002551	48	R-HEMBA1002551	869
	HEMBA1002767	F-HEMBA1002767	49	R-HEMBA1002767	870
	HEMBA1002985	F-HEMBA1002985	50	R-HEMBA1002985	871
	HEMBA1002992	F-HEMBA1002992	51		
10	HEMBA1003047	F-HEMBA1003047	52	R-HEMBA1003047	872
	HEMBA1003072	F-HEMBA1003072	53	R-HEMBA1003072	873
	HEMBA1003101	F-HEMBA1003101	54	R-HEMBA1003101	874
	HEMBA1003120	F-HEMBA1003120	55	R-HEMBA1003120	875
	HEMBA1003230	F-HEMBA1003230	56	R-HEMBA1003230	876
15	HEMBA1003294	F-HEMBA1003294	57	R-HEMBA1003294	877
	HEMBA1003315	F-HEMBA1003315	58	R-HEMBA1003315	878
	HEMBA1003392	F-HEMBA1003392	59	R-HEMBA1003392	879
	HEMBA1003399	F-HEMBA1003399	60	R-HEMBA1003399	880
20	HEMBA1003487	F-HEMBA1003487	61	R-HEMBA1003487	881
	HEMBA1003497	F-HEMBA1003497	62	R-HEMBA1003497	882
	HEMBA1003530	F-HEMBA1003530	63	R-HEMBA1003530	883
	HEMBA1003602	F-HEMBA1003602	64	R-HEMBA1003602	884
	HEMBA1003732	F-HEMBA1003732	65	R-HEMBA1003732	885
25	HEMBA1003945	F-HEMBA1003945	66	R-HEMBA1003945	886
	HEMBA1004007	F-HEMBA1004007	67	R-HEMBA1004007	887
	HEMBA1004067	F-HEMBA1004067	68		
	HEMBA1004085	F-HEMBA1004085	69	R-HEMBA1004085	888
30	HEMBA1004110	F-HEMBA1004110	70	R-nnnnnnnnnnnnn	889
	HEMBA1004250	F-HEMBA1004250	71	R-HEMBA1004250	890
	HEMBA1004391	F-HEMBA1004391	72	R-HEMBA1004391	891
	HEMBA1004444	F-HEMBA1004444	73	R-HEMBA1004444	892
	HEMBA1004454	F-HEMBA1004454	74	R-HEMBA1004454	893
35	HEMBA1004505	F-HEMBA1004505	75	R-HEMBA1004505	894
	HEMBA1004785	F-HEMBA1004785	76	R-HEMBA1004785	895
	HEMBA1004797	F-HEMBA1004797	77	R-HEMBA1004797	896
	HEMBA1004952	F-HEMBA1004952	78	R-HEMBA1004952	897
	HEMBA1004971	F-HEMBA1004971	79	R-HEMBA1004971	898
40	HEMBA1004982	F-HEMBA1004982	80	R-HEMBA1004982	899
	HEMBA1005070	F-HEMBA1005070	81	R-HEMBA1005070	900
	HEMBA1005084	F-HEMBA1005084	82	R-HEMBA1005084	901
	HEMBA1005145	F-HEMBA1005145	83	R-HEMBA1005145	902
45	HEMBA1005230	F-HEMBA1005230	84	R-HEMBA1005230	903
	HEMBA1005246	F-HEMBA1005246	85	R-HEMBA1005246	904
	HEMBA1005267	F-HEMBA1005267	86	R-HEMBA1005267	905
	HEMBA1005337	F-HEMBA1005337	87	R-HEMBA1005337	906
	HEMBA1005430	F-HEMBA1005430	88	R-HEMBA1005430	907
50	HEMBA1005449	F-HEMBA1005449	89	R-HEMBA1005449	908
	HEMBA1005489	F-HEMBA1005489	90	R-HEMBA1005489	909
	HEMBA1005522	F-HEMBA1005522	91	R-HEMBA1005522	910
	HEMBA1005545	F-HEMBA1005545	92	R-HEMBA1005545	911
55	HEMBA1005698	F-HEMBA1005698	93	R-HEMBA1005698	912
	HEMBA1005913	F-HEMBA1005913	94	R-HEMBA1005913	913
	HEMBA1005929	F-HEMBA1005929	95	R-HEMBA1005929	914

EP 1 130 094 A2

Table 1 (continued)

Correspondence between names of clone and the sequence name, and the SEQ ID.					
	Name of clone	Name of 5'-sequence	SEQ ID	3 Name of 3'-sequence	SEQ ID
5	HEMBA1005945	F-HEMBA1005945	96	R-HEMBA1005945	915
	HEMBA1006016	F-HEMBA1006016	97	R-HEMBA1006016	916
	HEMBA1006171	F-HEMBA1006171	98	R-HEMBA1006171	917
	HEMBA1006276	F-HEMBA1006276	99	R-HEMBA1006276	918
10	HEMBA1006299	F-HEMBA1006299	100	R-HEMBA1006299	919
	HEMBA1006311	F-HEMBA1006311	101	R-HEMBA1006311	920
	HEMBA1006335	F-HEMBA1006335	102	R-HEMBA1006335	921
	HEMBA1006357	F-HEMBA1006357	103	R-HEMBA1006357	922
	HEMBA1006430	F-HEMBA1006430	104	R-HEMBA1006430	923
15	HEMBA1006482	F-HEMBA1006482	105	R-HEMBA1006482	924
	HEMBA1006517	F-HEMBA1006517	106	R-HEMBA1006517	925
	HEMBA1006544	F-HEMBA1006544	107	R-HEMBA1006544	926
	HEMBA1006572	F-HEMBA1006572	108	R-HEMBA1006572	927
20	HEMBA1006658	F-HEMBA1006658	109	R-HEMBA1006658	928
	HEMBA1006707	F-HEMBA1006707	110	R-HEMBA1006707	929
	HEMBA1006724	F-HEMBA1006724	111	R-HEMBA1006724	930
	HEMBA1006749	F-HEMBA1006749	112	R-HEMBA1006749	931
	HEMBA1006770	F-HEMBA1006770	113	R-HEMBA1006770	932
25	HEMBA1006902	F-HEMBA1006902	114	R-HEMBA1006902	933
	HEMBA1006912	F-HEMBA1006912	115	R-HEMBA1006912	934
	HEMBA1006916	F-HEMBA1006916	116	R-HEMBA1006916	935
	HEMBA1006960	F-HEMBA1006960	117	R-HEMBA1006960	936
30	HEMBA1007013	F-HEMBA1007013	118	R-HEMBA1007013	937
	HEMBA1007057	F-HEMBA1007057	119	R-HEMBA1007057	938
	HEMBA1007063	F-HEMBA1007063	120	R-HEMBA1007063	939
	HEMBA1007226	F-HEMBA1007226	121		
	HEMBA1007241	F-HEMBA1007241	122	R-HEMBA1007241	940
35	HEMBA1007291	F-HEMBA1007291	123	R-HEMBA1007291	941
	HEMBA1007332	F-HEMBA1007332	124	R-HEMBA1007332	942
	HEMBA1000106	F-HEMBA1000106	125	R-HEMBA1000106	943
	HEMBA1000276	F-HEMBA1000276	126	R-HEMBA1000276	944
40	HEMBA1000309	F-HEMBA1000309	127	R-HEMBA1000309	945
	HEMBA1000407	F-HEMBA1000407	128	R-HEMBA1000407	946
	HEMBA1000447	F-HEMBA1000447	129	R-HEMBA1000447	947
	HEMBA1000542	F-HEMBA1000542	130	R-HEMBA1000542	948
	HEMBA1000567	F-HEMBA1000567	131	R-HEMBA1000567	949
45	HEMBA1000642	F-HEMBA1000642	132	R-HEMBA1000642	950
	HEMBA1000668	F-HEMBA1000668	133	R-HEMBA1000668	951
	HEMBA1000679	F-HEMBA1000679	134	R-HEMBA1000679	952
	HEMBA1000881	F-HEMBA1000881	135	R-HEMBA1000881	953
	HEMBA1000905	F-HEMBA1000905	136	R-HEMBA1000905	954
50	HEMBA1001026	F-HEMBA1001026	137	R-HEMBA1001026	955
	HEMBA1001048	F-HEMBA1001048	138	R-HEMBA1001048	956
	HEMBA1001200	F-HEMBA1001200	139	R-HEMBA1001200	957
	HEMBA1001407	F-HEMBA1001407	140	R-HEMBA1001407	958
55	HEMBA1001530	F-HEMBA1001530	141	R-HEMBA1001530	959
	HEMBA1001547	F-HEMBA1001547	142	R-HEMBA1001547	960
	HEMBA1001573	F-HEMBA1001573	143	R-HEMBA1001573	961

EP 1 130 094 A2

Table 1 (continued)

Correspondence between names of clone and the sequence name, and the SEQ ID.					
	Name of clone	Name of 5'-sequence	SEQ ID	3 Name of 3'-sequence	SEQ ID
5	HEMBB1001847	F-HEMBB1001847	144	R-HEMBB1001847	962
	HEMBB1001959	F-HEMBB1001959	145	R-HEMBB1001959	963
	HEMBB1001978	F-HEMBB1001978	146	R-HEMBB1001978	964
	HEMBB1002039	F-HEMBB1002039	147	R-HEMBB1002039	965
10	HEMBB1002041	F-HEMBB1002041	148	R-HEMBB1002041	966
	HEMBB1002051	F-HEMBB1002051	149	R-HEMBB1002051	967
	HEMBB1002120	F-HEMBB1002120	150	R-HEMBB1002120	968
	HEMBB1002162	F-HEMBB1002162	151	R-HEMBB1002162	969
15	HEMBB1002228	F-HEMBB1002228	152	R-HEMBB1002228	970
	HEMBB1002245	F-HEMBB1002245	153	R-HEMBB1002245	971
	HEMBB1002302	F-HEMBB1002302	154	R-HEMBB1002302	972
	HEMBB1002427	F-HEMBB1002427	155	R-HEMBB1002427	973
20	HEMBB1002465	F-HEMBB1002465	156	R-HEMBB1002465	974
	HEMBB1002661	F-HEMBB1002661	157	R-HEMBB1002661	975
	HEMBB1002663	F-HEMBB1002663	158	R-HEMBB1002663	976
	HEMBB1002693	F-HEMBB1002693	159	R-HEMBB1002693	977
25	MAMMA 1000046	F-MAMMA1000046	160	R-MAMMA1000046	978
	MAMMA1000102	F-MAMMA1000102	161	R-MAMMA1000102	979
	MAMMA 1000106	F-MAMMA1000106	162	R-MAMMA1000106	980
	MAMMA1000118	F-MAMMA1000118	163	R-MAMMA1000118	981
30	MAMMA1000141	F-MAMMA1000141	164	R-MAMMA1000141	982
	MAMMA1000204	F-MAMMA1000204	165	R-MAMMA1000204	983
	MAMMA1000226	F-MAMMA1000226	166	R-MAMMA1000226	984
	MAMMA1000403	F-MAMMA1000403	167	R-MAMMA1000403	985
35	MAMMA1000449	F-MAMMA1000449	168	R-MAMMA1000449	986
	MAMMA1000457	F-MAMMA1000457	169	R-MAMMA1000457	987
	MAMMA1000473	F-MAMMA1000473	170	R-MAMMA1000473	988
	MAMMA1000496	F-MAMMA1000496	171	R-MAMMA1000496	989
40	MAMMA1000528	F-MAMMA1000528	172	R-MAMMA1000528	990
	MAMMA1000591	F-MAMMA1000591	173	R-MAMMA1000591	991
	MAMMA 1000614	F-MAMMA1000614	174	R-MAMMA1000614	992
	MAMMA1000652	F-MAMMA1000652	175	R-MAMMA1000652	993
45	MAMMA1000681	F-MAMMA1000681	176	R-MAMMA1000681	994
	MAMMA 1000706	F-MAMMA1000706	177	R-MAMMA1000706	995
	MAMMA 1000788	F-MAMMA1000788	178	R-MAMMA1000788	996
	MAMMA 1000810	F-MAMMA1000810	179	R-MAMMA1000810	997
50	MAMMA1000814	F-MAMMA1000814	180	R-MAMMA1000814	998
	MAMMA1000881	F-MAMMA1000881	181	R-MAMMA1000881	999
	MAMMA1000986	F-MAMMA1000986	182	R-MAMMA1000986	1000
	MAMMA 1000994	F-MAMMA1000994	183	R-MAMMA1000994	1001
55	MAMMA1001043	F-MAMMA1001043	184	R-MAMMA1001043	1002
	MAMMA1001066	F-MAMMA1001066	185	R-MAMMA1001066	1003
	MAMMA1001094	F-MAMMA1001094	186	R-MAMMA1001094	1004
	MAMMA1001141	F-MAMMA1001141	187	R-MAMMA1001141	1005
	MAMMA1001150	F-MAMMA1001150	188	R-MAMMA1001150	1006
	MAMMA1001237	F-MAMMA1001237	189	R-MAMMA1001237	1007
	MAMMA1001284	F-MAMMA1001284	190	R-MAMMA1001284	1008
	MAMMA1001310	F-MAMMA1001310	191	R-MAMMA1001310	1009

EP 1 130 094 A2

Table 1 (continued)

Correspondence between names of clone and the sequence name, and the SEQ ID.					
	Name of clone	Name of 5'-sequence	SEQ ID	3 Name of 3'-sequence	SEQ ID
5	MAMMA1001344	F-MAMMA1001344	192		
	MAMMA1001418	F-MAMMA1001418	193	R-MAMMA1001418	1010
	MAMMA1001532	F-MAMMA1001532	194	R-MAMMA1001532	1011
	MAMMA1001609	F-MAMMA1001609	195	R-MAMMA1001609	1012
10	MAMMA 1001615	F-MAMMA1001615	196	R-MAMMA1001615	1013
	MAMMA 1001623	F-MAMMA1001623	197	R-MAMMA1001623	1014
	MAMMA1001634	F-MAMMA1001634	198	R-MAMMA1001634	1015
	MAMMA 1001893	F-MAMMA1001893	199	R-MAMMA1001893	1016
	MAMMA1001901	F-MAMMA1001901	200	R-MAMMA1001901	1017
15	MAMMA 1001957	F-MAMMA1001957	201	R-MAMMA1001957	1018
	MAMMA1001978	F-MAMMA1001978	202	R-MAMMA1001978	1019
	MAMMA1002070	F-MAMMA1002070	203	R-MAMMA1002070	1020
	MAMMA1002080	F-MAMMA1002080	204	R-MAMMA1002080	1021
20	MAMMA 1002087	F-MAMMA1002087	205	R-MAMMA1002087	1022
	MAMMA1002091	F-MAMMA1002091	206		
	MAMMA 1002095	F-MAMMA1002095	207	R-MAMMA1002095	1023
	MAMMA 1002128	F-MAMMA1002128	208	R-MAMMA1002128	1024
	MAMMA1002142	F-MAMMA1002142	209	R-MAMMA1002142	1025
25	MAMMA1002165	F-MAMMA1002165	210	R-MAMMA1002165	1026
	MAMMA 1002205	F-MAMMA1002205	211	R-MAMMA1002205	1027
	MAMMA 1002224	F-MAMMA1002224	212	R-MAMMA1002224	1028
	MAMMA1002234	F-MAMMA1002234	213	R-MAMMA1002234	1029
30	MAMMA1002586	F-MAMMA1002586	214	R-MAMMA1002586	1030
	MAMMA1002633	F-MAMMA1002633	215	R-MAMMA1002633	1031
	MAMMA1003126	F-MAMMA1003126	216	R-MAMMA1003126	1032
	NT2RM1000407	F-NT2RM1000407	217		
	NT2RM1000462	F-NT2RM1000462	218		
35	NT2RM1000542	F-NT2RM1000542	219		
	NT2RM1000580	F-NT2RM1000580	220		
	NT2RM1000789	F-NT2RM1000789	221		
	NT2RM1000855	F-NT2RM1000855	222		
40	NT2RM1000858	F-NT2RM1000858	223		
	NT2RM1000899	F-NT2RM1000899	224		
	NT2RM2000241	F-NT2RM2000241	225		
	NT2RM2000306	F-NT2RM2000306	226		
45	NT2RM2000410	F-NT2RM2000410	227		
	NT2RM2000423	F-NT2RM2000423	228		
	NT2RM2000497	F-NT2RM2000497	229		
	NT2RM2000514	F-NT2RM2000514	230		
	NT2RM2000565	F-NT2RM2000565	231		
50	NT2RM2000582	F-NT2RM2000582	232		
	NT2RM2000589	F-NT2RM2000589	233		
	NT2RM2000622	F-NT2RM2000622	234		
	NT2RM2000632	F-NT2RM2000632	235		
	NT2RM2000773	F-NT2RM2000773	236		
55	NT2RM2001126	F-NT2RM2001126	237		
	NT2RM2001558	F-NT2RM2001558	238		
	NT2RM2001626	F-NT2RM2001626	239		

EP 1 130 094 A2

Table 1 (continued)

Correspondence between names of clone and the sequence name, and the SEQ ID.					
	Name of clone	Name of 5'-sequence	SEQ ID	3 Name of 3'-sequence	SEQ ID
5	NT2RM2001643	F-NT2RM2001643	240		
	NT2RM2001738	F-NT2RM2001738	241		
	NT2RM2001767	F-NT2RM2001767	242		
	NT2RM2001792	F-NT2RM2001792	243		
10	NT2RM2001818	F-NT2RM2001818	244		
	NT2RM2001902	F-NT2RM2001902	245		
	NT2RM2001939	F-NT2RM2001939	246		
	NT2RM2001941	F-NT2RM2001941	247		
	NT2RM4000100	F-NT2RM4000100	248	R-NT2RM4000100	1033
15	NT2RM4000115	F-NT2RM4000115	249	R-NT2RM4000115	1034
	NT2RM4000198	F-NT2RM4000198	250	R-NT2RM4000198	1035
	NT2RM4000284	F-NT2RM4000284	251	R-NT2RM4000284	1036
	NT2RM4000295	F-NT2RM4000295	252	R-NT2RM4000295	1037
20	NT2RM4000326	F-NT2RM4000326	253	R-NT2RM4000326	1038
	NT2RM4000417	F-NT2RM4000417	254	R-NT2RM4000417	1039
	NT2RM4000444	F-NT2RM4000444	255	R-NT2RM4000444	1040
	NT2RM4000587	F-NT2RM4000587	256	R-NT2RM4000587	1041
	NT2RM4000593	F-NT2RM4000593	257	R-NT2RM4000593	1042
25	NT2RM4000648	F-NT2RM4000648	258	R-NT2RM4000648	1043
	NT2RM4000761	F-NT2RM4000761	259	R-NT2RM4000761	1044
	NT2RM4000965	F-NT2RM4000965	260	R-NT2RM4000965	1045
	NT2RM4000997	F-NT2RM4000997	261	R-NT2RM4000997	1046
30	NT2RM4001321	F-NT2RM4001321	262	R-NT2RM4001321	1047
	NT2RM4001325	F-NT2RM4001325	263	R-NT2RM4001325	1048
	NT2RM4001377	F-NT2RM4001377	264	R-NT2RM4001377	1049
	NT2RM4001735	F-NT2RM4001735	265	R-NT2RM4001735	1050
	NT2RM4001768	F-NT2RM4001768	266	R-NT2RM4001768	1051
35	NT2RM4001843	F-NT2RM4001843	267	R-NT2RM4001843	1052
	NT2RM4002352	F-NT2RM4002352	268	R-NT2RM4002352	1053
	NT2RP1000002	F-NT2RP1000002	269		
	NT2RP1000050	F-NT2RP1000050	270		
40	NT2RP1000181	F-NT2RP1000181	271		
	NT2RP1000239	F-NT2RP1000239	272		
	NT2RP1000261	F-NT2RP1000261	273		
	NT2RP1000271	F-NT2RP1000271	274		
	NT2RP1000300	F-NT2RP1000300	275		
45	NT2RP1000325	F-NT2RP1000325	276		
	NT2RP1000448	F-NT2RP1000448	277		
	NT2RP1000465	F-NT2RP1000465	278		
	NT2RP1000468	F-NT2RP1000468	279		
	NT2RP1000551	F-NT2RP1000551	280		
50	NT2RP1000579	F-NT2RP1000579	281		
	NT2RP1000613	F-NT2RP1000613	282		
	NT2RP1000679	F-NT2RP1000679	283		
	NT2RP1000740	F-NT2RP1000740	284		
55	NT2RP1000903	F-NT2RP1000903	285		
	NT2RP1000981	F-NT2RP1000981	286		
	NT2RP1001004	F-NT2RP1001004	287		

EP 1 130 094 A2

Table 1 (continued)

Correspondence between names of clone and the sequence name, and the SEQ ID.					
	Name of clone	Name of 5'-sequence	SEQ ID	3 Name of 3'-sequence	SEQ ID
5	NT2RP1001020	F-NT2RP1001020	288		
	NT2RP1001031	F-NT2RP1001031	289		
	NT2RP1001563	F-NT2RP1001563	290		
	NT2RP2000092	F-NT2RP2000092	291	R-NT2RP2000092	1054
10	NT2RP2000178	F-NT2RP2000178	292	R-NT2RP2000178	1055
	NT2RP2000240	F-NT2RP2000240	293	R-NT2RP2000240	1056
	NT2RP2000394	F-NT2RP2000394	294	R-NT2RP2000394	1057
	NT2RP2000447	F-NT2RP2000447	295	R-NT2RP2000447	1058
	NT2RP2000479	F-NT2RP2000479	296	R-NT2RP2000479	1059
15	NT2RP2000514	F-NT2RP2000514	297	R-NT2RP2000514	1060
	NT2RP2000533	F-NT2RP2000533	298	R-NT2RP2000533	1061
	NT2RP2000610	F-NT2RP2000610	299		
	NT2RP2000616	F-NT2RP2000616	300	R-NT2RP2000616	1062
20	NT2RP2000649	F-NT2RP2000649	301	R-NT2RP2000649	1063
	NT2RP2000663	F-NT2RP2000663	302	R-NT2RP2000663	1064
	NT2RP2000694	F-NT2RP2000694	303		
	NT2RP2000712	F-NT2RP2000712	304	R-NT2RP2000712	1065
	NT2RP2000739	F-NT2RP2000739	305	R-NT2RP2000739	1066
25	NT2RP2000818	F-NT2RP2000818	306	R-NT2RP2000818	1067
	NT2RP2000903	F-NT2RP2000903	307	R-NT2RP2000903	1068
	NT2RP2001200	F-NT2RP2001200	308	R-NT2RP2001200	1069
	NT2RP2001223	F-NT2RP2001223	309	R-NT2RP2001223	1070
30	NT2RP2001276	F-NT2RP2001276	310	R-NT2RP2001276	1071
	NT2RP2001388	F-NT2RP2001388	311	R-NT2RP2001388	1072
	NT2RP2001469	F-NT2RP2001469	312	R-NT2RP2001469	1073
	NT2RP2001480	F-NT2RP2001480	313	R-NT2RP2001480	1074
	NT2RP2001495	F-NT2RP2001495	314	R-NT2RP2001495	1075
35	NT2RP2001514	F-NT2RP2001514	315	R-NT2RP2001514	1076
	NT2RP2001529	F-NT2RP2001529	316		
	NT2RP2001538	F-NT2RP2001538	317	R-NT2RP2001538	1077
	NT2RP2001562	F-NT2RP2001562	318	R-NT2RP2001562	1078
40	NT2RP2001662	F-NT2RP2001662	319	R-NT2RP2001662	1079
	NT2RP2001755	F-NT2RP2001755	320	R-NT2RP2001755	1080
	NT2RP2001769	F-NT2RP2001769	321	R-NT2RP2001769	1081
	NT2RP2001817	F-NT2RP2001817	322	R-NT2RP2001817	1082
	NT2RP2001878	F-NT2RP2001878	323	R-NT2RP2001878	1083
45	NT2RP2001903	F-NT2RP2001903	324	R-NT2RP2001903	1084
	NT2RP2001915	F-NT2RP2001915	325	R-NT2RP2001915	1085
	NT2RP2001921	F-NT2RP2001921	326	R-NT2RP2001921	1086
	NT2RP2001948	F-NT2RP2001948	327	R-NT2RP2001948	1087
50	NT2RP2001956	F-NT2RP2001956	328	R-NT2RP2001956	1088
	NT2RP2002015	F-NT2RP2002015	329	R-NT2RP2002015	1089
	NT2RP2002063	F-NT2RP2002063	330	R-NT2RP2002063	1090
	NT2RP2002188	F-NT2RP2002188	331	R-NT2RP2002188	1091
	NT2RP2002232	F-NT2RP2002232	332	R-NT2RP2002232	1092
55	NT2RP2002304	F-NT2RP2002304	333	R-nnnnnnnnnnnnn	1093
	NT2RP2002409	F-NT2RP2002409	334	R-NT2RP2002409	1094
	NT2RP2002510	F-NT2RP2002510	335	R-NT2RP2002510	1095

EP 1 130 094 A2

Table 1 (continued)

Correspondence between names of clone and the sequence name, and the SEQ ID.					
	Name of clone	Name of 5'-sequence	SEQ ID	3 Name of 3'-sequence	SEQ ID
5	NT2RP2002527	F-NT2RP2002527	336	R-NT2RP2002527	1096
	NT2RP2002533	F-NT2RP2002533	337	R-NT2RP2002533	1097
	NT2RP2002564	F-NT2RP2002564	338	R-NT2RP2002564	1098
	NT2RP2002674	F-NT2RP2002674	339	R-NT2RP2002674	1099
10	NT2RP2002721	F-NT2RP2002721	340	R-NT2RP2002721	1100
	NT2RP2002824	F-NT2RP2002824	341	R-NT2RP2002824	1101
	NT2RP2002942	F-NT2RP2002942	342	R-NT2RP2002942	1102
	NT2RP2002974	F-NT2RP2002974	343	R-NT2RP2002974	1103
	NT2RP2002976	F-NT2RP2002976	344	R-NT2RP2002976	1104
15	NT2RP2003042	F-NT2RP2003042	345	R-NT2RP2003042	1105
	NT2RP2003138	F-NT2RP2003138	346		
	NT2RP2003179	F-NT2RP2003179	347	R-NT2RP2003179	1106
	NT2RP2003210	F-NT2RP2003210	348	R-NT2RP2003210	1107
20	NT2RP2003302	F-NT2RP2003302	349	R-NT2RP2003302	1108
	NT2RP2003369	F-NT2RP2003369	350	R-NT2RP2003369	1109
	NT2RP2003383	F-NT2RP2003383	351	R-NT2RP2003383	1110
	NT2RP2003390	F-NT2RP2003390	352	R-NT2RP2003390	1111
	NT2RP2003469	F-NT2RP2003469	353	R-NT2RP2003469	1112
25	NT2RP2003545	F-NT2RP2003545	354	R-NT2RP2003545	1113
	NT2RP2003593	F-NT2RP2003593	355	R-NT2RP2003593	1114
	NT2RP2003599	F-NT2RP2003599	356	R-NT2RP2003599	1115
	NT2RP2003655	F-NT2RP2003655	357	R-NT2RP2003655	1116
30	NT2RP2003664	F-NT2RP2003664	358	R-NT2RP2003664	1117
	NT2RP2003931	F-NT2RP2003931	359	R-NT2RP2003931	1118
	NT2RP2003940	F-NT2RP2003940	360	R-NT2RP2003940	1119
	NT2RP2003950	F-NT2RP2003950	361	R-NT2RP2003950	1120
	NT2RP2004069	F-NT2RP2004069	362	R-NT2RP2004069	1121
35	NT2RP2004108	F-NT2RP2004108	363	R-NT2RP2004108	1122
	NT2RP2004141	F-NT2RP2004141	364	R-NT2RP2004141	1123
	NT2RP2004179	F-NT2RP2004179	365	R-NT2RP2004179	1124
	NT2RP2004205	F-NT2RP2004205	366	R-NT2RP2004205	1125
40	NT2RP2004447	F-NT2RP2004447	367	R-NT2RP2004447	1126
	NT2RP2004495	F-NT2RP2004495	368	R-NT2RP2004495	1127
	NT2RP2004524	F-NT2RP2004524	369	R-NT2RP2004524	1128
	NT2RP2004556	F-NT2RP2004556	370	R-NT2RP2004556	1129
	NT2RP2004606	F-NT2RP2004606	371	R-NT2RP2004606	1130
45	NT2RP2004648	F-NT2RP2004648	372	R-NT2RP2004648	1131
	NT2RP2004670	F-NT2RP2004670	373	R-NT2RP2004670	1132
	NT2RP2004794	F-NT2RP2004794	374	R-NT2RP2004794	1133
	NT2RP2004837	F-NT2RP2004837	375	R-NT2RP2004837	1134
	NT2RP2004847	F-NT2RP2004847	376	R-NT2RP2004847	1135
50	NT2RP2005027	F-NT2RP2005027	377	R-NT2RP2005027	1136
	NT2RP2005069	F-NT2RP2005069	378	R-NT2RP2005069	1137
	NT2RP2005163	F-NT2RP2005163	379	R-NT2RP2005163	1138
	NT2RP2005181	F-NT2RP2005181	380	R-NT2RP2005181	1139
55	NT2RP2005247	F-NT2RP2005247	381	R-NT2RP2005247	1140
	NT2RP2005378	F-NT2RP2005378	382	R-NT2RP2005378	1141
	NT2RP2005391	F-NT2RP2005391	383	R-NT2RP2005391	1142

EP 1 130 094 A2

Table 1 (continued)

Correspondence between names of clone and the sequence name, and the SEQ ID.					
	Name of clone	Name of 5'-sequence	SEQ ID	3 Name of 3'-sequence	SEQ ID
5	NT2RP2005425	F-NT2RP2005425	384	R-NT2RP2005425	1143
	NT2RP2005463	F-NT2RP2005463	385	R-NT2RP2005463	1144
	NT2RP2005514	F-NT2RP2005514	386	R-NT2RP2005514	1145
	NT2RP2005535	F-NT2RP2005535	387	R-NT2RP2005535	1146
10	NT2RP2005541	F-NT2RP2005541	388	R-NT2RP2005541	1147
	NT2RP2005597	F-NT2RP2005597	389	R-NT2RP2005597	1148
	NT2RP2005632	F-NT2RP2005632	390	R-ntntntntntntntntntntnt	1149
	NT2RP2005666	F-NT2RP2005666	391	R-NT2RP2005666	1150
	NT2RP2005774	F-NT2RP2005774	392	R-NT2RP2005774	1151
15	NT2RP2005878	F-NT2RP2005878	393	R-NT2RP2005878	1152
	NT2RP2005883	F-NT2RP2005883	394	R-NT2RP2005883	1153
	NT2RP2005887	F-NT2RP2005887	395	R-NT2RP2005887	1154
	NT2RP2005941	F-NT2RP2005941	396	R-ntntntntntntntntntntnt	1155
20	NT2RP2005994	F-NT2RP2005994	397	R-NT2RP2005994	1156
	NT2RP2006004	F-NT2RP2006004	398	R-NT2RP2006004	1157
	NT2RP2006042	F-NT2RP2006042	399	R-NT2RP2006042	1158
	NT2RP2006092	F-NT2RP2006092	400	R-NT2RP2006092	1159
	NT2RP2006099	F-NT2RP2006099	401	R-NT2RP2006099	1160
25	NT2RP2006134	F-NT2RP2006134	402	R-NT2RP2006134	1161
	NT2RP2006269	F-NT2RP2006269	403	R-NT2RP2006269	1162
	NT2RP2006512	F-NT2RP2006512	404	R-NT2RP2006512	1163
	NT2RP3000011	F-NT2RP3000011	405	R-NT2RP3000011	1164
	NT2RP3000022	F-NT2RP3000022	406	R-NT2RP3000022	1165
30	NT2RP3000059	F-NT2RP3000059	407	R-NT2RP3000059	1166
	NT2RP3000063	F-NT2RP3000063	408	R-NT2RP3000063	1167
	NT2RP3000125	F-NT2RP3000125	409	R-ntntntntntntntntntntnt	1168
	NT2RP3000148	F-NT2RP3000148	410	R-NT2RP3000148	1169
35	NT2RP3000169	F-NT2RP3000169	411	R-NT2RP3000169	1170
	NT2RP3000171	F-NT2RP3000171	412	R-NT2RP3000171	1171
	NT2RP3000172	F-NT2RP3000172	413	R-NT2RP3000172	1172
	NT2RP3000201	F-NT2RP3000201	414	R-NT2RP3000201	1173
	NT2RP3000232	F-NT2RP3000232	415	R-NT2RP3000232	1174
40	NT2RP3000304	F-NT2RP3000304	416	R-NT2RP3000304	1175
	NT2RP3000378	F-NT2RP3000378	417	R-NT2RP3000378	1176
	NT2RP3000427	F-NT2RP3000427	418		
	NT2RP3000436	F-NT2RP3000436	419	R-NT2RP3000436	1177
45	NT2RP3000444	F-NT2RP3000444	420	R-NT2RP3000444	1178
	NT2RP3000460	F-NT2RP3000460	421	R-NT2RP3000460	1179
	NT2RP3000481	F-NT2RP3000481	422	R-NT2RP3000481	1180
	NT2RP3000616	F-NT2RP3000616	423	R-NT2RP3000616	1181
	NT2RP3000645	F-NT2RP3000645	424	R-NT2RP3000645	1182
50	NT2RP3000652	F-NT2RP3000652	425	R-NT2RP3000652	1183
	NT2RP3000676	F-NT2RP3000676	426	R-NT2RP3000676	1184
	NT2RP3000677	F-NT2RP3000677	427	R-NT2RP3000677	1185
	NT2RP3000721	F-NT2RP3000721	428	R-NT2RP3000721	1186
55	NT2RP3000789	F-NT2RP3000789	429	R-NT2RP3000789	1187
	NT2RP3000818	F-NT2RP3000818	430	R-NT2RP3000818	1188
	NT2RP3000820	F-NT2RP3000820	431	R-NT2RP3000820	1189

EP 1 130 094 A2

Table 1 (continued)

Correspondence between names of clone and the sequence name, and the SEQ ID.					
	Name of clone	Name of 5'-sequence	SEQ ID	3 Name of 3'-sequence	SEQ ID
5	NT2RP3000838	F-NT2RP3000838	432	R-NT2RP3000838	1190
	NT2RP3000871	F-NT2RP3000871	433	R-NT2RP3000871	1191
	NT2RP3000907	F-NT2RP3000907	434	R-NT2RP3000907	1192
	NT2RP3000921	F-NT2RP3000921	435	R-NT2RP3000921	1193
10	NT2RP3001012	F-NT2RP3001012	436	R-NT2RP3001012	1194
	NT2RP3001044	F-NT2RP3001044	437	R-NT2RP3001044	1195
	NT2RP3001061	F-NT2RP3001061	438	R-NT2RP3001061	1196
	NT2RP3001159	F-NT2RP3001159	439	R-NT2RP3001159	1197
	NT2RP3001170	F-NT2RP3001170	440	R-NT2RP3001170	1198
15	NT2RP3001195	F-NT2RP3001195	441	R-NT2RP3001195	1199
	NT2RP3001240	F-NT2RP3001240	442	R-NT2RP3001240	1200
	NT2RP3001271	F-NT2RP3001271	443	R-NT2RP3001271	1201
	NT2RP3001322	F-NT2RP3001322	444	R-NT2RP3001322	1202
20	NT2RP3001388	F-NT2RP3001388	445		
	NT2RP3001542	F-NT2RP3001542	446	R-NT2RP3001542	1203
	NT2RP3001560	F-NT2RP3001560	447	R-NT2RP3001560	1204
	NT2RP3001592	F-NT2RP3001592	448	R-NT2RP3001592	1205
	NT2RP3001650	F-NT2RP3001650	449		
25	NT2RP3001685	F-NT2RP3001685	450	R-NT2RP3001685	1206
	NT2RP3001738	F-NT2RP3001738	451	R-NT2RP3001738	1207
	NT2RP3001754	F-NT2RP3001754	452	R-NT2RP3001754	1208
	NT2RP3001858	F-NT2RP3001858	453	R-NT2RP3001858	1209
30	NT2RP3001976	F-NT2RP3001976	454	R-NT2RP3001976	1210
	NT2RP3002015	F-NT2RP3002015	455	R-NT2RP3002015	1211
	NT2RP3002160	F-NT2RP3002160	456	R-NT2RP3002160	1212
	NT2RP3002281	F-NT2RP3002281	457	R-NT2RP3002281	1213
	NT2RP3002286	F-NT2RP3002286	458	R-NT2RP3002286	1214
35	NT2RP3002311	F-NT2RP3002311	459	R-NT2RP3002311	1215
	NT2RP3002324	F-NT2RP3002324	460	R-NT2RP3002324	1216
	NT2RP3002342	F-NT2RP3002342	461	R-NT2RP3002342	1217
	NT2RP3002353	F-NT2RP3002353	462	R-NT2RP3002353	1218
40	NT2RP3002409	F-NT2RP3002409	463	R-NT2RP3002409	1219
	NT2RP3002411	F-NT2RP3002411	464	R-NT2RP3002411	1220
	NT2RP3002448	F-NT2RP3002448	465	R-NT2RP3002448	1221
	NT2RP3002571	F-NT2RP3002571	466	R-NT2RP3002571	1222
	NT2RP3002664	F-NT2RP3002664	467	R-NT2RP3002664	1223
45	NT2RP3002721	F-NT2RP3002721	468	R-NT2RP3002721	1224
	NT2RP3002737	F-NT2RP3002737	469	R-NT2RP3002737	1225
	NT2RP3002738	F-NT2RP3002738	470	R-NT2RP3002738	1226
	NT2RP3002790	F-NT2RP3002790	471	R-NT2RP3002790	1227
	NT2RP3002836	F-NT2RP3002836	472	R-NT2RP3002836	1228
50	NT2RP3002887	F-NT2RP3002887	473	R-NT2RP3002887	1229
	NT2RP3002900	F-NT2RP3002900	474	R-NT2RP3002900	1230
	NT2RP3002958	F-NT2RP3002958	475	R-NT2RP3002958	1231
	NT2RP3002983	F-NT2RP3002983	476	R-NT2RP3002983	1232
55	NT2RP3003000	F-NT2RP3003000	477	R-NT2RP3003000	1233
	NT2RP3003076	F-NT2RP3003076	478	R-NT2RP3003076	1234
	NT2RP3003354	F-NT2RP3003354	479	R-NT2RP3003354	1235

EP 1 130 094 A2

Table 1 (continued)

Correspondence between names of clone and the sequence name, and the SEQ ID.					
	Name of clone	Name of 5'-sequence	SEQ ID	3 Name of 3'-sequence	SEQ ID
5	NT2RP3003448	F-NT2RP3003448	480	R-NT2RP3003448	1236
	NT2RP3003469	F-NT2RP3003469	481	R-NT2RP3003469	1237
	NT2RP3003473	F-NT2RP3003473	482	R-NT2RP3003473	1238
	NT2RP3003527	F-NT2RP3003527	483	R-NT2RP3003527	1239
10	NT2RP3003532	F-NT2RP3003532	484	R-NT2RP3003532	1240
	NT2RP3003535	F-NT2RP3003535	485	R-rrrrrrrrrrrrrrrrrr	1241
	NT2RP3003559	F-NT2RP3003559	486	R-NT2RP3003559	1242
	NT2RP3003614	F-NT2RP3003614	487	R-NT2RP3003614	1243
15	NT2RP3003729	F-NT2RP3003729	488	R-NT2RP3003729	1244
	NT2RP3003849	F-NT2RP3003849	489	R-NT2RP3003849	1245
	NT2RP3003874	F-NT2RP3003874	490	R-NT2RP3003874	1246
	NT2RP3003939	F-NT2RP3003939	491		
20	NT2RP3003963	F-NT2RP3003963	492	R-NT2RP3003963	1247
	NT2RP3004000	F-NT2RP3004000	493	R-NT2RP3004000	1248
	NT2RP3004025	F-NT2RP3004025	494	R-NT2RP3004025	1249
	NT2RP3004067	F-NT2RP3004067	495		
25	NT2RP3004075	F-NT2RP3004075	496	R-NT2RP3004075	1250
	NT2RP3004083	F-NT2RP3004083	497	R-NT2RP3004083	1251
	NT2RP3004090	F-NT2RP3004090	498	R-NT2RP3004090	1252
	NT2RP3004119	F-NT2RP3004119	499	R-NT2RP3004119	1253
30	NT2RP3004130	F-NT2RP3004130	500	R-NT2RP3004130	1254
	NT2RP3004133	F-NT2RP3004133	501	R-NT2RP3004133	1255
	NT2RP3004202	F-NT2RP3004202	502	R-NT2RP3004202	1256
	NT2RP3004294	F-NT2RP3004294	503	R-NT2RP3004294	1257
35	NT2RP3004309	F-NT2RP3004309	504	R-NT2RP3004309	1258
	NT2RP3004321	F-NT2RP3004321	505	R-NT2RP3004321	1259
	NT2RP3004345	F-NT2RP3004345	506	R-NT2RP3004345	1260
	NT2RP3004355	F-NT2RP3004355	507	R-NT2RP3004355	1261
40	NT2RP3004374	F-NT2RP3004374	508	R-NT2RP3004374	1262
	NT2RP3004406	F-NT2RP3004406	509	R-NT2RP3004406	1263
	NT2RP3004481	F-NT2RP3004481	510	R-NT2RP3004481	1264
	NT2RP3004552	F-NT2RP3004552	511	R-NT2RP3004552	1265
45	NT2RP3004557	F-NT2RP3004557	512		
	NT2RP3004625	F-NT2RP3004625	513	R-NT2RP3004625	1266
	NT2RP3004640	F-NT2RP3004640	514	R-NT2RP3004640	1267
	NT2RP3004647	F-NT2RP3004647	515	R-NT2RP3004647	1268
50	NT2RP4000108	F-NT2RP4000108	516	R-NT2RP4000108	1269
	NT2RP4000634	F-NT2RP4000634	517	R-NT2RP4000634	1270
	NT2RP4000962	F-NT2RP4000962	518	R-NT2RP4000962	1271
	NT2RP4001001	F-NT2RP4001001	519	R-NT2RP4001001	1272
55	NT2RP4001009	F-NT2RP4001009	520	R-NT2RP4001009	1273
	NT2RP4001467	F-NT2RP4001467	521	R-NT2RP4001467	1274
	NT2RP4001877	F-NT2RP4001877	522	R-NT2RP4001877	1275
	NT2RP4001879	F-NT2RP4001879	523	R-NT2RP4001879	1276
55	NT2RP4002187	F-NT2RP4002187	524	R-NT2RP4002187	1277
	NT2RP4002451	F-NT2RP4002451	525	R-NT2RP4002451	1278
	NT2RP4002715	F-NT2RP4002715	526	R-NT2RP4002715	1279
	NT2RP4002750	F-NT2RP4002750	527	R-NT2RP4002750	1280

EP 1 130 094 A2

Table 1 (continued)

Correspondence between names of clone and the sequence name, and the SEQ ID.					
	Name of clone	Name of 5'-sequence	SEQ ID	3 Name of 3'-sequence	SEQ ID
5	OVARC1000003	F-OVARC1000003	528	R-OVARC1000003	1281
	OVARC1000090	F-OVARC1000090	529	R-OVARC1000090	1282
	OVARC1000105	F-OVARC1000105	530	R-OVARC1000105	1283
	OVARC1000137	F-OVARC1000137	531	R-OVARC1000137	1284
10	OVARC1000208	F-OVARC1000208	532	R-OVARC1000208	1285
	OVARC1000255	F-OVARC1000255	533	R-OVARC1000255	1286
	OVARC1000275	F-OVARC1000275	534	R-OVARC1000275	1287
	OVARC1000298	F-OVARC1000298	535	R-OVARC1000298	1288
	OVARC1000307	F-OVARC1000307	536	R-OVARC1000307	1289
15	OVARC1000313	F-OVARC1000313	537	R-OVARC1000313	1290
	OVARC1000331	F-OVARC1000331	538	R-OVARC1000331	1291
	OVARC1000410	F-OVARC1000410	539	R-OVARC1000410	1292
	OVARC1000439	F-OVARC1000439	540	R-OVARC1000439	1293
20	OVARC1000467	F-OVARC1000467	541	R-OVARC1000467	1294
	OVARC1000529	F-OVARC1000529	542	R-OVARC1000529	1295
	OVARC1000553	F-OVARC1000553	543	R-OVARC1000553	1296
	OVARC1000775	F-OVARC1000775	544	R-OVARC1000775	1297
	OVARC1000811	F-OVARC1000811	545	R-OVARC1000811	1298
25	OVARC1000853	F-OVARC1000853	546	R-OVARC1000853	1299
	OVARC1000873	F-OVARC1000873	547	R-OVARC1000873	1300
	OVARC1000916	F-OVARC1000916	548	R-OVARC1000916	1301
	OVARC1000956	F-OVARC1000956	549	R-OVARC1000956	1302
30	OVARC1000995	F-OVARC1000995	550	R-OVARC1000995	1303
	OVARC1001030	F-OVARC1001030	551	R-OVARC1001030	1304
	OVARC1001049	F-OVARC1001049	552	R-OVARC1001049	1305
	OVARC1001086	F-OVARC1001086	553	R-OVARC1001086	1306
	OVARC1001132	F-OVARC1001132	554	R-OVARC1001132	1307
35	OVARC1001163	F-OVARC1001163	555	R-OVARC1001163	1308
	OVARC1001222	F-OVARC1001222	556	R-OVARC1001222	1309
	OVARC1001260	F-OVARC1001260	557	R-OVARC1001260	1310
	OVARC1001336	F-OVARC1001336	558	R-OVARC1001336	1311
40	OVARC1001338	F-OVARC1001338	559	R-OVARC1001338	1312
	OVARC1001569	F-OVARC1001569	560	R-OVARC1001569	1313
	OVARC1001570	F-OVARC1001570	561	R-OVARC1001570	1314
	OVARC1001596	F-OVARC1001596	562	R-OVARC1001596	1315
	OVARC1001607	F-OVARC1001607	563	R-OVARC1001607	1316
45	OVARC1001725	F-OVARC1001725	564	R-OVARC1001725	1317
	OVARC1001727	F-OVARC1001727	565	R-OVARC1001727	1318
	OVARC1001807	F-OVARC1001807	566	R-OVARC1001807	1319
	OVARC1001833	F-OVARC1001833	567	R-OVARC1001833	1320
	OVARC1001952	F-OVARC1001952	568		
50	OVARC1001991	F-OVARC1001991	569	R-OVARC1001991	1321
	OVARC1002058	F-OVARC1002058	570	R-OVARC1002058	1322
	OVARC1002178	F-OVARC1002178	571	R-OVARC1002178	1323
	PLACE 1000033	F-PLACE1000033	572	R-PLACE1000033	1324
55	PLACE 1000231	F-PLACE1000231	573	R-PLACE1000231	1325
	PLACE 1000258	F-PLACE1000258	574	R-PLACE1000258	1326
	PLACE 1000442	F-PLACE1000442	575	R-PLACE1000442	1327

EP 1 130 094 A2

Table 1 (continued)

Correspondence between names of clone and the sequence name, and the SEQ ID.					
	Name of clone	Name of 5'-sequence	SEQ ID	3 Name of 3'-sequence	SEQ ID
5	PLACE 1000560	F-PLACE1000560	576	R-PLACE1000560	1328
	PLACE 1000740	F-PLACE1000740	577	R-PLACE1000740	1329
	PLACE 1000907	F-PLACE1000907	578		
10	PLACE 1000912	F-PLACE1000912	579	R-PLACE1000912	1330
	PLACE 1000914	F-PLACE1000914	580	R-PLACE1000914	1331
	PLACE 1000927	F-PLACE1000927	581	R-PLACE1000927	1332
	PLACE 1000986	F-PLACE1000986	582	R-PLACE1000986	1333
	PLACE1001016	F-PLACE1001016	583	R-PLACE1001016	1334
15	PLACE 1001100	F-PLACE1001100	584	R-PLACE1001100	1335
	PLACE 1001114	F-PLACE1001114	585	R-PLACE1001114	1336
	PLACE 1001123	F-PLACE1001123	586	R-PLACE1001123	1337
	PLACE 1001183	F-PLACE1001183	587	R-PLACE1001183	1338
	PLACE 1001229	F-PLACE1001229	588	R-PLACE1001229	1339
20	PLACE 1001231	F-PLACE1001231	589	R-PLACE1001231	1340
	PLACE 1001340	F-PLACE1001340	590	R-PLACE1001340	1341
	PLACE1001401	F-PLACE1001401	591	R-PLACE1001401	1342
	PLACE 1001407	F-PLACE1001407	592	R-PLACE1001407	1343
	PLACE 1001464	F-PLACE1001464	593	R-PLACE1001464	1344
25	PLACE1001500	F-PLACE1001500	594	R-PLACE1001500	1345
	PLACE1001516	F-PLACE1001516	595	R-PLACE1001516	1346
	PLACE1001536	F-PLACE1001536	596	R-PLACE1001536	1347
	PLACE 1001564	F-PLACE1001564	597	R-PLACE1001564	1348
	PLACE 1001655	F-PLACE1001655	598	R-PLACE1001655	1349
30	PLACE 1001788	F-PLACE1001788	599	R-PLACE1001788	1350
	PLACE1001795	F-PLACE1001795	600	R-PLACE1001795	1351
	PLACE1001836	F-PLACE1001836	601	R-PLACE1001836	1352
	PLACE 1001918	F-PLACE1001918	602	R-PLACE1001918	1353
	PLACE 1001949	F-PLACE1001949	603	R-PLACE1001949	1354
35	PLACE 1002080	F-PLACE1002080	604	R-PLACE1002080	1355
	PLACE 1002095	F-PLACE1002095	605	R-PLACE1002095	1356
	PLACE 1002153	F-PLACE1002153	606	R-PLACE1002153	1357
	PLACE 1002329	F-PLACE1002329	607	R-PLACE1002329	1358
	PLACE 1002355	F-PLACE1002355	608	R-PLACE1002355	1359
40	PLACE1002374	F-PLACE1002374	609	R-PLACE1002374	1360
	PLACE 1002518	F-PLACE1002518	610	R-PLACE1002518	1361
	PLACE 1002547	F-PLACE1002547	611	R-PLACE1002547	1362
	PLACE 1002726	F-PLACE1002726	612	R-PLACE1002726	1363
	PLACE 1002905	F-PLACE1002905	613	R-PLACE1002905	1364
45	PLACE 1002911	F-PLACE1002911	614	R-PLACE1002911	1365
	PLACE 1002967	F-PLACE1002967	615	R-PLACE1002967	1366
	PLACE 1003135	F-PLACE1003135	616	R-PLACE1003135	1367
	PLACE1003163	F-PLACE1003163	617	R-PLACE1003163	1368
	PLACE 1003407	F-PLACE1003407	618	R-PLACE1003407	1369
50	PLACE 1003428	F-PLACE1003428	619	R-PLACE1003428	1370
	PLACE1003438	F-PLACE1003438	620	R-PLACE1003438	1371
	PLACE 1003460	F-PLACE1003460	621	R-PLACE1003460	1372
	PLACE 1003529	F-PLACE1003529	622	R-nnnnnnnnnnnnn	1373
	PLACE 1003573	F-PLACE1003573	623	R-PLACE1003573	1374

EP 1 130 094 A2

Table 1 (continued)

Correspondence between names of clone and the sequence name, and the SEQ ID.					
	Name of clone	Name of 5'-sequence	SEQ ID	3 Name of 3'-sequence	SEQ ID
5	PLACE 1003598	F-PLACE1003598	624	R-PLACE1003598	1375
	PLACE1003644	F-PLACE1003644	625	R-PLACE1003644	1376
	PLACE1003737	F-PLACE1003737	626	R-PLACE1003737	1377
	PLACE 1003772	F-PLACE1003772	627	R-PLACE1003772	1378
10	PLACE 1003839	F-PLACE1003839	628	R-PLACE1003839	1379
	PLACE 1003845	F-PLACE1003845	629	R-PLACE1003845	1380
	PLACE 1003 852	F-PLACE1003852	630	R-PLACE1003852	1381
	PLACE 1004028	F-PLACE1004028	631	R-PLACE1004028	1382
15	PLACE 1004078	F-PLACE1004078	632	R-PLACE1004078	1383
	PLACE1004166	F-PLACE1004166	633	R-PLACE1004166	1384
	PLACE 1004168	F-PLACE1004168	634	R-nnnnnnnnnnnnn	1385
	PLACE1004199	F-PLACE1004199	635	R-PLACE1004199	1386
20	PLACE 1004279	F-PLACE1004279	636	R-PLACE1004279	1387
	PLACE 1004282	F-PLACE1004282	637	R-PLACE1004282	1388
	PLACE 1004305	F-PLACE1004305	638	R-PLACE1004305	1389
	PLACE 1004441	F-PLACE1004441	639	R-PLACE1004441	1390
25	PLACE 1004450	F-PLACE1004450	640	R-PLACE1004450	1391
	PLACE 1004482	F-PLACE1004482	641	R-PLACE1004482	1392
	PLACE 1004492	F-PLACE1004492	642	R-PLACE1004492	1393
	PLACE 1004519	F-PLACE1004519	643	R-PLACE1004519	1394
30	PLACE 1004520	F-PLACE1004520	644	R-PLACE1004520	1395
	PLACE 1004630	F-PLACE1004630	645	R-PLACE1004630	1396
	PLACE 1004637	F-PLACE1004637	646	R-PLACE1004637	1397
	PLACE 1004648	F-PLACE1004648	647	R-PLACE1004648	1398
35	PLACE 1004816	F-PLACE1004816	648	R-PLACE1004816	1399
	PLACE 1004887	F-PLACE1004887	649	R-PLACE1004887	1400
	PLACE 1005003	F-PLACE1005003	650	R-PLACE1005003	1401
	PLACE 1005005	F-PLACE1005005	651	R-PLACE1005005	1402
40	PLACE 1005031	F-PLACE1005031	652	R-PLACE1005031	1403
	PLACE 1005239	F-PLACE1005239	653	R-PLACE1005239	1404
	PLACE 1005250	F-PLACE1005250	654	R-PLACE1005250	1405
	PLACE 1005383	F-PLACE1005383	655	R-PLACE1005383	1406
45	PLACE 1005410	F-PLACE1005410	656	R-PLACE1005410	1407
	PLACE 1005426	F-PLACE1005426	657	R-PLACE1005426	1408
	PLACE 1005519	F-PLACE1005519	658	R-PLACE1005519	1409
	PLACE 1005539	F-PLACE1005539	659	R-PLACE1005539	1410
50	PLACE 1005544	F-PLACE1005544	660	R-PLACE1005544	1411
	PLACE 1005569	F-PLACE1005569	661	R-PLACE1005569	1412
	PLACE 1005601	F-PLACE1005601	662	R-PLACE1005601	1413
	PLACE 1005660	F-PLACE1005660	663	R-PLACE1005660	1414
55	PLACE 1005669	F-PLACE1005669	664	R-PLACE1005669	1415
	PLACE 1005682	F-PLACE1005682	665	R-PLACE1005682	1416
	PLACE 1005725	F-PLACE1005725	666	R-PLACE1005725	1417
	PLACE 1005736	F-PLACE1005736	667	R-PLACE1005736	1418
	PLACE 1005745	F-PLACE1005745	668	R-PLACE1005745	1419
	PLACE 1005768	F-PLACE1005768	669	R-PLACE1005768	1420
	PLACE1005815	F-PLACE1005815	670	R-PLACE1005815	1421
	PLACE 1005878	F-PLACE1005878	671	R-PLACE1005878	1422

EP 1 130 094 A2

Table 1 (continued)

Correspondence between names of clone and the sequence name, and the SEQ ID.					
	Name of clone	Name of 5'-sequence	SEQ ID	3 Name of 3'-sequence	SEQ ID
5	PLACE 1005927	F-PLACE1005927	672	R-PLACE1005927	1423
	PLACE 1006071	F-PLACE1006071	673	R-PLACE1006071	1424
	PLACE 1006073	F-PLACE1006073	674	R-PLACE1006073	1425
	PLACE 1006079	F-PLACE1006079	675	R-PLACE1006079	1426
10	PLACE 1006093	F-PLACE1006093	676	R-PLACE1006093	1427
	PLACE 1006208	F-PLACE1006208	677	R-nnnnnnnnnnnnn	1428
	PLACE 1006219	F-PLACE1006219	678	R-PLACE1006219	1429
	PLACE 1006277	F-PLACE1006277	679	R-PLACE1006277	1430
	PLACE 1006290	F-PLACE1006290	680	R-PLACE1006290	1431
15	PLACE 1006443	F-PLACE1006443	681	R-PLACE1006443	1432
	PLACE 1006515	F-PLACE1006515	682	R-PLACE1006515	1433
	PLACE 1006716	F-PLACE1006716	683	R-PLACE1006716	1434
	PLACE 1006786	F-PLACE1006786	684	R-PLACE1006786	1435
20	PLACE 1006809	F-PLACE1006809	685	R-PLACE1006809	1436
	PLACE 1006959	F-PLACE1006959	686	R-PLACE1006959	1437
	PLACE 1007028	F-PLACE1007028	687	R-PLACE1007028	1438
	PLACE 1007040	F-PLACE1007040	688	R-PLACE1007040	1439
	PLACE 1007077	F-PLACE1007077	689	R-PLACE1007077	1440
25	PLACE 1007081	F-PLACE1007081	690	R-PLACE1007081	1441
	PLACE 1007096	F-PLACE1007096	691	R-PLACE1007096	1442
	PLACE 1007296	F-PLACE1007296	692	R-PLACE1007296	1443
	PLACE 1007591	F-PLACE1007591	693	R-PLACE1007591	1444
30	PLACE1007626	F-PLACE1007626	694	R-PLACE1007626	1445
	PLACE 1007702	F-PLACE1007702	695	R-PLACE1007702	1446
	PLACE1007845	F-PLACE1007845	696	R-PLACE1007845	1447
	PLACE1007881	F-PLACE1007881	697	R-PLACE1007881	1448
	PLACE 1007971	F-PLACE1007971	698	R-PLACE1007971	1449
35	PLACE 1008282	F-PLACE1008282	699	R-PLACE1008282	1450
	PLACE 1008297	F-PLACE1008297	700	R-PLACE1008297	1451
	PLACE 1008359	F-PLACE1008359	701	R-PLACE1008359	1452
	PLACE 1008469	F-PLACE1008469	702	R-PLACE1008469	1453
40	PLACE 1008549	F-PLACE1008549	703	R-PLACE1008549	1454
	PLACE 1008657	F-PLACE1008657	704	R-PLACE1008657	1455
	PLACE 1008716	F-PLACE1008716	705	R-PLACE1008716	1456
	PLACE 1008744	F-PLACE1008744	706	R-PLACE1008744	1457
	PLACE 1008984	F-PLACE1008984	707	R-PLACE1008984	1458
45	PLACE 1008985	F-PLACE1008985	708	R-PLACE1008985	1459
	PLACE 1009067	F-PLACE1009067	709	R-PLACE1009067	1460
	PLACE 1009196	F-PLACE1009196	710	R-PLACE1009196	1461
	PLACE 1009279	F-PLACE1009279	711	R-PLACE1009279	1462
	PLACE 1009527	F-PLACE1009527	712	R-PLACE1009527	1463
50	PLACE 1009546	F-PLACE1009546	713	R-PLACE1009546	1464
	PLACE 1009600	F-PLACE1009600	714	R-PLACE1009600	1465
	PLACE 1009735	F-PLACE1009735	715	R-PLACE1009735	1466
	PLACE 1009982	F-PLACE1009982	716	R-nnnnnnnnnnnnn	1467
55	PLACE1010011	F-PLACE1010011	717	R-PLACE1010011	1468
	PLACE1010078	F-PLACE1010078	718	R-PLACE1010078	1469
	PLACE1010081	F-PLACE1010081	719	R-PLACE1010081	1470

EP 1 130 094 A2

Table 1 (continued)

Correspondence between names of clone and the sequence name, and the SEQ ID.					
	Name of clone	Name of 5'-sequence	SEQ ID	3 Name of 3'-sequence	SEQ ID
5	PLACE 1010251	F-PLACE1010251	720	R-PLACE1010251	1471
	PLACE1010445	F-PLACE1010445	721	R-PLACE1010445	1472
	PLACE 1010713	F-PLACE1010713	722	R-PLACE1010713	1473
	PLACE 1010784	F-PLACE1010784	723	R-PLACE1010784	1474
10	PLACE 1010827	F-PLACE1010827	724	R-PLACE1010827	1475
	PLACE 1010968	F-PLACE1010968	725	R-PLACE1010968	1476
	PLACE 1011045	F-PLACE1011045	726	R-PLACE1011045	1477
	PLACE1011116	F-PLACE1011116	727	R-PLACE1011116	1478
	PLACE1011181	F-PLACE1011181	728		
15	PLACE1011236	F-PLACE1011236	729	R-PLACE1011236	1479
	PLACE 1011364	F-PLACE1011364	730	R-PLACE1011364	1480
	PLACE 1 011407	F-PLACE1011407	731	R-PLACE1011407	1481
	PLACE1011516	F-PLACE1011516	732	R-PLACE1011516	1482
20	PLACE 1011708	F-PLACE1011708	733	R-PLACE1011708	1483
	PLACE 1011824	F-PLACE1011824	734	R-PLACE1011824	1484
	PLACE 1011978	F-PLACE1011978	735	R-PLACE1011978	1485
	PLACE2000118	F-PLACE2000118	736	R-PLACE2000118	1486
	PLACE2000219	F-PLACE2000219	737	R-PLACE2000219	1487
25	PLACE3000181	F-PLACE3000181	738	R-PLACE3000181	1488
	PLACE3000213	F-PLACE3000213	739	R-PLACE3000213	1489
	PLACE4000354	F-PLACE4000354	740	R-PLACE4000354	1490
	PLACE4000455	F-PLACE4000455	741	R-PLACE4000455	1491
	SKNMC1000004	F-SKNMC1000004	742		
30	SKNMC1000014	F-SKNMC1000014	743		
	SKNMC1000082	F-SKNMC1000082	744		
	THYRO1000036	F-THYRO1000036	745	R-THYRO1000036	1492
	THYRO1000061	F-THYRO1000061	746	R-THYRO1000061	1493
35	THYRO1000099	F-THYRO1000099	747	R-THYRO1000099	1494
	THYRO1000196	F-THYRO1000196	748	R-THYRO1000196	1495
	THYRO1000400	F-THYRO1000400	749	R-THYRO1000400	1496
	THYRO1000580	F-THYRO1000580	750	R-THYRO1000580	1497
40	THYRO1000584	F-THYRO1000584	751	R-THYRO1000584	1498
	THYRO1000678	F-THYRO1000678	752	R-THYRO1000678	1499
	THYRO1000776	F-THYRO1000776	753	R-THYRO1000776	1500
	THYRO1000795	F-THYRO1000795	754	R-THYRO1000795	1501
	THYRO1000846	F-THYRO1000846	755	R-THYRO1000846	1502
45	THYRO1000866	F-THYRO1000866	756	R-THYRO1000866	1503
	THYRO1000956	F-THYRO1000956	757	R-THYRO1000956	1504
	THYRO1000964	F-THYRO1000964	758	R-THYRO1000964	1505
	THYRO1000999	F-THYRO1000999	759	R-THYRO1000999	1506
	THYRO1001063	F-THYRO1001063	760	R-THYRO1001063	1507
50	THYRO1001071	F-THYRO1001071	761	R-THYRO1001071	1508
	THYRO1001102	F-THYRO1001102	762	R-THYRO1001102	1509
	THYRO1001113	F-THYRO1001113	763	R-THYRO1001113	1510
	THYRO1001128	F-THYRO1001128	764	R-THYRO1001128	1511
55	THYRO1001205	F-THYRO1001205	765	R-THYRO1001205	1512
	THYRO1001237	F-THYRO1001237	766	R-THYRO1001237	1513
	THYRO1001242	F-THYRO1001242	767	R-THYRO1001242	1514

EP 1 130 094 A2

Table 1 (continued)

Correspondence between names of clone and the sequence name, and the SEQ ID.					
	Name of clone	Name of 5'-sequence	SEQ ID	3 Name of 3'-sequence	SEQ ID
5	THYRO1001266	F-THYRO1001266	768	R-THYRO1001266	1515
	THYRO1001327	F-THYRO1001327	769	R-THYRO1001327	1516
	THYRO1001456	F-THYRO1001456	770	R-THYRO1001456	1517
	THYRO1001457	F-THYRO1001457	771	R-THYRO1001457	1518
10	THYRO1001471	F-THYRO1001471	772	R-THYRO1001471	1519
	THYRO1001478	F-THYRO1001478	773	R-THYRO1001478	1520
	THYRO1001495	F-THYRO1001495	774	R-THYRO1001495	1521
	THYRO1001523	F-THYRO1001523	775	R-THYRO1001523	1522
	THYRO1001529	F-THYRO1001529	776	R-THYRO1001529	1523
15	THYRO1001593	F-THYRO1001593	777	R-THYRO1001593	1524
	THYRO1001608	F-THYRO1001608	778	R-THYRO1001608	1525
	THYRO1001641	F-THYRO1001641	779	R-THYRO1001641	1526
	THYRO1001700	F-THYRO1001700	780	R-THYRO1001700	1527
20	THYRO1001702	F-THYRO1001702	781	R-THYRO1001702	1528
	THYRO1001725	F-THYRO1001725	782	R-THYRO1001725	1529
	THYRO1001770	F-THYRO1001770	783	R-THYRO1001770	1530
	THYRO1001803	F-THYRO1001803	784	R-THYRO1001803	1531
	Y79AA1000030	F-Y79AA1000030	785	R-Y79AA1000030	1532
25	Y79AA1000127	F-Y79AA1000127	786	R-Y79AA1000127	1533
	Y79AA1000207	F-Y79AA1000207	787	R-Y79AA1000207	1534
	Y79AA1000226	F-Y79AA1000226	788	R-Y79AA1000226	1535
	Y79AA1000270	F-Y79AA1000270	789	R-Y79AA1000270	1536
30	Y79AA1000426	F-Y79AA1000426	790	R-Y79AA1000426	1537
	Y79AA1000521	F-Y79AA1000521	791	R-Y79AA1000521	1538
	Y79AA1000750	F-Y79AA1000750	792	R-Y79AA1000750	1539
	Y79AA1000776	F-Y79AA1000776	793	R-Y79AA1000776	1540
	Y79AA1000777	F-Y79AA1000777	794	R-Y79AA1000777	1541
35	Y79AA1000876	F-Y79AA1000876	795	R-Y79AA1000876	1542
	Y79AA1000888	F-Y79AA1000888	796		
	Y79AA1000959	F-Y79AA1000959	797	R-Y79AA1000959	1543
	Y79AA1000967	F-Y79AA1000967	798	R-Y79AA1000967	1544
40	Y79AA1001013	F-Y79AA1001013	799	R-Y79AA1001013	1545
	Y79AA1001056	F-Y79AA1001056	800	R-Y79AA1001056	1546
	Y79AA1001062	F-Y79AA1001062	801	R-Y79AA1001062	1547
	Y79AA1001090	F-Y79AA1001090	802	R-Y79AA1001090	1548
	Y79AA1001212	F-Y79AA1001212	803	R-Y79AA1001212	1549
45	Y79AA1001264	F-Y79AA1001264	804	R-Y79AA1001264	1550
	Y79AA1001272	F-Y79AA1001272	805	R-Y79AA1001272	1551
	Y79AA1001328	F-Y79AA1001328	806	R-Y79AA1001328	1552
	Y79AA1001426	F-Y79AA1001426	807	R-Y79AA1001426	1553
	Y79AA1001427	F-Y79AA1001427	808		
50	Y79AA1001430	F-Y79AA1001430	809	R-Y79AA1001430	1554
	Y79AA1001523	F-Y79AA1001523	810	R-Y79AA1001523	1555
	Y79AA1001530	F-Y79AA1001530	811	R-Y79AA1001530	1556
	Y79AA1001592	F-Y79AA1001592	812	R-Y79AA1001592	1557
55	Y79AA1001727	F-Y79AA1001727	813	R-Y79AA1001727	1558
	Y79AA1001787	F-Y79AA1001787	814	R-Y79AA1001787	1559
	Y79AA1001793	F-Y79AA1001793	815		

EP 1 130 094 A2

Table 1 (continued)

Correspondence between names of clone and the sequence name, and the SEQ ID.				
Name of clone	Name of 5'-sequence	SEQ ID	3 Name of 3'-sequence	SEQ ID
Y79AA1001795	F-Y79AA1001795	816	R-Y79AA1001795	1560
Y79AA1001799	F-Y79AA1001799	817	R-Y79AA1001799	1561
Y79AA1001803	F-Y79AA1001803	818	R-Y79AA1001803	1562
Y79AA1001863	F-Y79AA1001863	819	R-Y79AA1001863	1563
Y79AA1002022	F-Y79AA1002022	820	R-Y79AA1002022	1564
Y79AA1002058	F-Y79AA1002058	821		
Y79AA1002121	F-Y79AA1002121	822	R-nnnnnnnnnnnn	1565
Y79AA1002129	F-Y79AA1002129	823	R-nnnnnnnnnnnn	1566
Y79AA1002213	F-Y79AA1002213	824	R-Y79AA1002213	1567
Y79AA1002334	F-Y79AA1002334	825	R-Y79AA1002334	1568
Y79AA1002373	F-Y79AA1002373	826	R-Y79AA1002373	1569
Y79AA1002376	F-Y79AA1002376	827	R-Y79AA1002376	1570
Y79AA1002378	F-Y79AA1002378	828	R-Y79AA1002378	1571
Y79AA1002381	F-Y79AA1002381	829	R-Y79AA1002381	1572
NT2RP2006580	F-NT2RP2006580	2545	R-NT2RP2006580	2546

The sequence name starting from "F" means the name of 5'-end sequence, and the sequence name starting from "R" means the name of 3'-end sequence. A blank indicates that the 3'-end sequence corresponding to the 5'-end sequence has not been determined in the clone.

[0018] Furthermore, the present invention relates to the use of the above primers, as described below.

- (4) A polynucleotide which can be synthesized with the primer set of (2) or (3).
- (5) A polynucleotide comprising a coding region in the polynucleotide of (4).
- (6) A substantially pure protein encoded by polynucleotide of (4).
- (7) A partial peptide of the protein of (6).

[0019] In addition, the present invention comprises a polynucleotide described below and a protein encoded by the polynucleotide.

- (8) An isolated polynucleotide selected from the group consisting of

- (a) a polynucleotide comprising a coding region of the nucleotide sequence set forth in any one of the SEQ ID NOs in Table 370;
- (b) a polynucleotide comprising a nucleotide sequence encoding a protein comprising the amino acid sequence set forth in any one of the SEQ ID NOs in Table 370;
- (c) a polynucleotide comprising a nucleotide sequence encoding a protein comprising an amino acid sequence selected from the amino acid sequences set forth in the SEQ ID NOs in Table 370, in which one or more amino acids are substituted, deleted, inserted, and/or added, wherein said protein is functionally equivalent to the protein comprising said amino acid sequence selected from the amino acid sequences set forth in the SEQ ID NOs in Table 370;
- (d) a polynucleotide that hybridizes with a polynucleotide comprising a nucleotide sequence selected from the nucleotide sequences set forth in the SEQ ID NOs in Table 370, and that comprises a nucleotide sequence encoding a protein functionally equivalent to the protein encoded by the nucleotide sequence selected from the nucleotide sequences set forth in the SEQ ID NOs in Table 370;
- (e) a polynucleotide comprising a nucleotide sequence encoding a partial amino acid sequence of a protein encoded by the polynucleotide of (a) to (d);
- (f) a polynucleotide comprising a nucleotide sequence with at least 70% identity to the nucleotide sequence set forth in any one of the SEQ ID NOs in Table 370.

- (9). A substantially pure protein encoded by the polynucleotide of (8).
- (10) An antibody against the protein or peptide of any one of (6), (7), and (9).
- (11) A vector comprising the polynucleotide of (5) or (8).

- (12) A transformant carrying the polynucleotide of (5) or (8), or the vector of (11).
 (13) A transformant expressively carrying the polynucleotide of (5) or (8), or the vector of (11).
 (14) A method for producing the protein or peptide of any one of (6), (7), and (9), comprising culturing the transformant of (13) and recovering the expression product.
 5 (15) An oligonucleotide comprising; the nucleotide sequence set forth in any one of the SEQ ID NOs in Table 370 or the nucleotide sequence complementary to the complementary strand thereof, wherein said oligonucleotide comprises 15 nucleotides or more.
 (16) Use of the oligonucleotide of (15) as a primer for synthesizing a polynucleotide.
 (17) Use of the oligonucleotide of (15) as a probe for detecting a gene.
 10 (18) An antisense polynucleotide against the polynucleotide of (8), or the portion thereof.
 (19) A method for synthesizing a polynucleotide, the method comprising:
 a) synthesizing a complementary strand using a cDNA library as a template, and using the primer set of (2) or (3), or the primer of (16); and
 15 b) recovering the synthesized product.
 (20) The method of (19), wherein the cDNA library is obtainable by oligo-capping method.
 (21) The method of (19), wherein the complementary strand is obtainable by PCR.
 20 (22) A method for detecting the polynucleotide of (8), the method comprising:
 a) incubating a target polynucleotide with the oligonucleotide of (15) under the conditions where hybridization occurs, and
 b) detecting the hybridization of the target polynucleotide with the oligonucleotide of (15).
 25 (23) A database of polynucleotides and/or proteins, the database comprising information on at least one sequence selected from the nucleotide sequences set forth in the SEQ ID NOs in Table 370 and/or the amino acid sequences set forth in the SEQ ID NOs in Table 370, or a medium on which the database is stored.

[0020] Any patents, patent applications, and publications cited herein are incorporated by reference.

BRIEF DESCRIPTION OF THE DRAWINGS

[0021]

35 Figure 1 shows the restriction maps of vectors pME18SFL3 and pUC19FL3.
 Figure 2 shows the reproducibility of gene expression analysis. The ordinate and the abscissa show the intensities of gene expression obtained in experiments different from each other.
 Figure 3 shows the detection limit in gene expression analysis. The intensity of expression is shown in the ordinate, and the concentration ($\mu\text{g/ml}$) of the probe used is shown in the abscissa.

DETAILED DESCRIPTION OF THE INVENTION

[0022] Herein, "polynucleotide" is defined as a molecule in which multiple nucleotides are polymerized. There are no limitations in the number of the polymerized nucleotides. In case that the polymer contains relatively low number of nucleotides, it is also described as an "oligonucleotide". The polynucleotide or the oligonucleotide of the present invention can be a natural or chemically synthesized product. Alternatively, it can be synthesized using a template DNA by an enzymatic reaction such as PCR.

[0023] All the cDNA provided by the invention are full-length cDNA. Herein, a "full-length cDNA" is defined as a cDNA which contains both ATG codon (the translation start site) and the stop codon. Accordingly, the untranslated regions, which are originally found in the upstream or downstream of the protein coding region in natural mRNA, may or may not be contained.

An "isolated polynucleotide" is a polynucleotide the structure of which is not identical to that of any naturally occurring nucleic acid or to that of any fragment of a naturally occurring genomic nucleic acid spanning more than three separate genes. The term therefore covers, for example,

- 55 (a) a DNA which has the sequence of part of a naturally occurring genomic DNA molecule but is not flanked by both of the coding sequences that flank that part of the molecule in the genome of the organism in which it naturally occurs;

(b) a nucleic acid incorporated into a vector or into the genomic DNA of a prokaryote or eukaryote in a manner such that the resulting molecule is not identical to any naturally occurring vector or genomic DNA;

(c) a separate molecule such as a cDNA, a genomic fragment, a fragment produced by polymerase chain reaction (PCR), or a restriction fragment; and

(d) a recombinant nucleotide sequence that is part of a hybrid gene, i.e., a gene encoding a fusion protein. Specifically excluded from this definition are nucleic acids present in mixtures of different (i) DNA molecules, (ii) transfected cells, or (iii) cell clones: e.g., as these occur in a DNA library such as a cDNA or genomic DNA library.

[0024] The term "substantially pure" as used herein in reference to a given polypeptide means that the protein or polypeptide is substantially free from other biological macromolecules. The substantially pure protein or polypeptide is at least 75% (e.g., at least 80, 85, 95, or 99%) pure by dry weight. Purity can be measured by any appropriate standard method, for example, by column chromatography, polyacrylamide gel electrophoresis, or HPLC analysis.

[0025] All the clones of the present invention (830 clones) are novel and covering full-length, and also predicted to encode any of the following functional protein:

secretory proteins,
membrane proteins,
proteins associated to signal transduction (signal transduction-associated proteins; e.g. protein kinases, etc.),
proteins associated to a glycoprotein (glycoprotein-associated proteins),
proteins associated with transcription (transcription-associated proteins),
proteins associated with diseases (disease-associated proteins),
or, enzymes and/or metabolism-associated proteins, cell division- and/or cell proliferation-associated proteins,
cytoskeleton-associated proteins, nuclear proteins, DNA-and/or RNA-binding proteins, ATP- and/or GTP-binding proteins, protein synthesis- and/or protein transport-associated proteins, and cellular defense-associated proteins.

[0026] Furthermore, all the cDNA clones of the present invention can be characterized as follows:

(1) a cDNA that is obtained by the oligo-capping method, which provides cDNA with high fullness ratio. The cDNA was selected by the score in the ATGpr (described as ATGpr1, as well), which is a program for prediction of the fullness of the 5'-end of cDNA based on the features of the 5'-end sequence. In addition, the PSORT, which is a program for prediction of the existence of the signal sequence selected, cDNA that contains a signal sequence in the 5'-end, or transmembrane region in the protein coding region. Furthermore, the homology search with the 5'-end sequences confirmed that, the selected clones were not identical to any of the known human mRNA (namely novel);

or,
(2) a cDNA that is obtained by the oligo-capping method, which provides cDNA with high fullness ratio. The cDNA was selected by the score in the ATGpr, which is a program for prediction of the fullness of the 5'-end based on the features of the 5'-end sequence. Furthermore, the a cDNA that has relative homology with an amino acid sequence of a protein with known functions was selected by the BLAST search (Altschul S.F., Gish W., Miller W., Myers E.W., and Lipman D.J. (1990) J. Mol. Biol. 215: 403-410 ; Gish W., and States D.J. (1993) Nature Genet. 3: 266-272) on the SwissProt database using the 5'-end sequence. In addition, the homology search using the 5'-end sequence confirmed that the selected clones were not identical to any of the known human mRNA (namely novel).

[0027] All clones are obtainable as a full-length clone by such a method as PCR (Current Protocols in Molecular Biology, Ausubel et al. edit, (1987) John Wiley & Sons, Section 6.1-6.4) using both the 5'- and 3'-end sequences, or using the 5'-end sequence and an oligo-dT primer that corresponds to the polyA sequence.

[0028] Specifically, PCR can be performed using an oligonucleotide that has 15 nucleotides longer, and specifically hybridizes with the complementary strand of the polynucleotide that contains the nucleotide sequence selected from the 5'-end sequences shown in Table 1 (SEQ ID NO: 1-829, and SEQ ID NO: 2545), and an oligo-dT primer as a 5'-, and 3'-primer, respectively. The length of the primers is usually 15-100 bp, and favorably between 15-35 bp. In case of LA PCR, which is described below, the primer length of 25-35 bp may provide a good result.

[0029] A method to design a primer that enables a specific amplification based on the given nucleotide sequence is known to those skilled in the art (Current Protocols in Molecular Biology, Ausubel et al. edit, (1987) John Wiley & Sons, Section 6.1-6.4). In designing a primer based on the 5'-end sequence, the primer is designed so as that, in principle, the amplification products will include the translation start site. Accordingly, in case that a given 5'-end nucleotide sequence is the 5'-untranslated region (5'UTR), any part of the sequence can be used as a 5'-primer as far as the specificity toward the target cDNA is insured. The translation start site can be predicted using a known method such

as the ATGpr as described below.

[0030] When synthesizing a polynucleotide, the target nucleotide sequence to be amplified can extend to several thousand bp in some cDNA. However, it is possible to amplify such a long nucleotides by using such as LA PCR (Long and Accurate PCR). It is advantageous to use LA PCR when synthesizing long DNA. In LA PCR, in which a special DNA polymerase having 3' \square 5' exonuclease activity is used, misincorporated nucleotides can be removed. Accordingly, accurate synthesis of the complementary strand can be achieved even with a long nucleotide sequence. By using LA PCR, it is reported that amplification of a nucleotide with 20 kb longer can be achieved under desirable condition (Takeshi Hayashi (1996) Jikken-Igaku Bessatsu, "Advanced Technologies in PCR" Youdo-sha).

[0031] A template DNA for synthesizing the cDNA of the present invention can be obtained by using cDNA libraries that are prepared by various methods. The full-length cDNA clones obtained here are those with high fullness ratio, which were obtained using a combination of (1) a method to prepare a full-length-enriched cDNA library using the oligo-capping method, and (2) an estimation system for fullness using the 5'-end sequence (selection based on the estimation by the ATGpr after removing clones that are non-full-length compared to the ESTs). However, it is possible to easily obtain a full-length cDNA by using the primers that are provided by the present invention, not by the above described specialized method.

[0032] The problem with the cDNA libraries prepared by the known methods or commercially available is that mRNA contained in the libraries has very low fullness ratio. Thus, it is difficult to screen full-length cDNA clone directly from the library using ordinary cloning methods. The present invention has revealed a primer that is capable of synthesizing a full-length cDNA. If provided with primers, it is possible to synthesize a target full-length cDNA by using enzymatic reactions such as PCR. In particular, a full-length-enriched cDNA library, synthesized by methods such as oligo-capping, is desirable to synthesize a full-length cDNA with more reliability.

[0033] Once the nucleotide sequences of the full-length cDNAs obtained in the present invention is determined, it is possible to predict the functions of the proteins encoded by the cDNA clones, for example, by searching the databases such as GenBank (<http://www.ncbi.nlm.nih.gov/web/GenBank/>), Swiss-Prot (http://www.ebi.ac.uk/ebi_docs/Swiss-Prot_db/swiss/home.html), UniGene (<http://www.ncbi.nlm.nih.gov/UniGene/>) for homologies of the cDNAs, or by searching the amino acid sequences deduced from the full-length nucleotide sequences for signal sequence by using software such as PSORT (K. Nakai & M. Kanehisa, Genomics, 14: 897-991 (1992), for transmembrane region by using software such as SOSUI (T. Hirokawa et al., Bioinformatics, 14:378-379 (1998); Mitsui Knowledge Industry Co., Ltd.) or for motif by using software such as Pfam (<http://www.sanger.ac.uk/Software/Pfam/index.shtml>) or PROSITE (<http://www.expasy.ch/prosite>). As a matter of course, the functions are often predictable by using partial sequence information (preferably 300 nucleotides or more) instead of the full-length nucleotide sequences. However, the result of the prediction obtained by using partial sequence information does not always agree with the result obtained by using full-length nucleotide sequence, and thus it is needless to say that the prediction of function is preferably performed based on the full-length nucleotide sequences.

[0034] Homology search using each of GenBank, Swiss-Prot and UniGene was performed for the 826 clones whose full-length nucleotide sequences had been determined (HEMBA1005337, NT2RM1000407, NT2RM2001767, and NT2RP3003939 are not full-length). The amino acid sequences deduced from the full-length nucleotide sequences were searched for functional domains by using analytical software programs, PSORT, SOSUI and Pfam. Based on the results, proteins encoded by the cDNA clones were grouped into some categories and their functions were predicted.

[0035] The following 437 clones were categorized into secretory and/or membrane proteins. The clones categorized into secretory and/or membrane proteins are those which matched the full-length sequences of Swiss-Prot database with the keywords "growth factor", "cytokine", "hormone", "signal", "transmembrane", "membrane", "extracellular matrix", "receptor", "G-protein coupled receptor", "ionic channel", "voltage-gated channel", "calcium channel", "cell adhesion", "collagen" or "connective tissue"; those which matched the data, suggesting that the proteins are secretory and/or membrane proteins; or those which matched with the full-length sequences of GenBank or UniGene database similar description; and, further, those predicted to have an N-terminal signal sequence or a transmembrane region as a result of domain search for the amino acid sequences deduced from the full-length nucleotide sequences.

BNGH41000020,	BNGH41000087,	BNGH41000091,	HEMBA1000121,	HEMBA1000128,	HEMBA1000349,
HEMBA1000477,	HEMBA1000590,	HEMBA1000713,	HEMBA1000732,	HEMBA1000745,	HEMBA1000835,
HEMBA1000940,	HEMBA1000962,	HEMBA1001221,	HEMBA1001228,	HEMBA1001621,	HEMBA1002131,
HEMBA1002163,	HEMBA1002167,	HEMBA1002178,	HEMBA1002195,	HEMBA1002227,	HEMBA1002420,
HEMBA1002421,	HEMBA1002767,	HEMBA1003047,	HEMBA1003101,	HEMBA1003230,	HEMBA1003392,
HEMBA1003530,	HEMBA1003602,	HEMBA1003732,	HEMBA1003945,	HEMBA1004110,	HEMBA1004250,
HEMBA1004391,	HEMBA1004444,	HEMBA1004454,	HEMBA1004505,	HEMBA1004797,	HEMBA1004982,
HEMBA1005070,	HEMBA1005449,	HEMBA1005522,	HEMBA1005545,	HEMBA1005698,	HEMBA1005945,
HEMBA1006171,	HEMBA1006299,	HEMBA1006311,	HEMBA1006335,	HEMBA1006357,	HEMBA1006430,
HEMBA1006482,	HEMBA1006707,	HEMBA1006724,	HEMBA1006749,	HEMBA1006902,	HEMBA1006960,
HEMBA1007241,	HEMBA1000407,	HEMBA1000447,	HEMBA1000567,	HEMBA1000679,	HEMBA1000881,

EP 1 130 094 A2

	HEMBB1001026,	HEMBB1001048,	HEMBB1001407,	HEMBB1001530,	HEMBB1001573,	HEMBB1001847,
	HEMBB1001978,	HEMBB1002041,	HEMBB1002162,	HEMBB1002245,	HEMBB1002427,	HEMBB1002693,
	MAMMA1000102,	MAMMA1000106,	MAMMA1000118,	MAMMA1000141,	MAMMA1000204,	MAMMA1000226,
	MAMMA1000457,	MAMMA1000473,	MAMMA1000496,	MAMMA1000591,	MAMMA1000681,	MAMMA1000810,
5	MAMMA1000986,	MAMMA1000994,	MAMMA1001043,	MAMMA1001141,	MAMMA1001237,	MAMMA1001344,
	MAMMA1001418,	MAMMA1001893,	MAMMA1001957,	MAMMA1001978,		
	MAMMA1002070,	MAMMA1002091,	MAMMA1002095,	MAMMA1002165,	MAMMA1002234,	MAMMA1002586,
	MAMMA1002633,	MAMMA1003126,	NT2RM1000462,	NT2RM1000542,	NT2RM1000580,	NT2RM1000855,
	NT2RM1000858,	NT2RM1000899,	NT2RM2000241,	NT2RM2000410,	NT2RM2000423,	NT2RM2000565,
10	NT2RM2001626,	NT2RM2001792,	NT2RM2001939,	NT2RM2001941,	NT2RM4000198,	NT2RM4000284,
	NT2RM4000417,	NT2RM4000444,	NT2RM4000587,	NT2RM4000593,	NT2RM4000648,	NT2RM4000761,
	NT2RM4000997,	NT2RM4001325,	NT2RM4001735,	NT2RM4001768,	NT2RM4001843,	NT2RM4002352,
	NT2RP1000050,	NT2RP1000181,	NT2RP1000261,	NT2RP1000300,	NT2RP1000325,	NT2RP1000448,
	NT2RP1000551,	NT2RP1000613,	NT2RP1000981,	NT2RP1001004,	NT2RP1001563,	NT2RP2000479,
15	NT2RP2000533,	NT2RP2000616,	NT2RP2000649,	NT2RP2000663,	NT2RP2000694,	NT2RP2000818,
	NT2RP2000903,	NT2RP2001200,	NT2RP2001276,	NT2RP2001480,	NT2RP2001495,	NT2RP2001514,
	NT2RP2001755,	NT2RP2001915,	NT2RP2001956,	NT2RP2002063,	NT2RP2002188,	NT2RP2002232,
	NT2RP2002527,	NT2RP2002533,	NT2RP2002721,	NT2RP2002824,	NT2RP2002942,	NT2RP2002976,
	NT2RP2003042,	NT2RP2003210,	NT2RP2003383,	NT2RP2003390,	NT2RP2003469,	NT2RP2003593,
20	NT2RP2003655,	NT2RP2003664,	NT2RP2003950,	NT2RP2004179,	NT2RP2004205,	NT2RP2004495,
	NT2RP2004524,	NT2RP2004556,	NT2RP2004606,	NT2RP2004648,	NT2RP2004794,	NT2RP2005027,
	NT2RP2005163,	NT2RP2005181,	NT2RP2005378,	NT2RP2005463,	NT2RP2005541,	NT2RP2005597,
	NT2RP2005666,	NT2RP2005883,	NT2RP2005994,	NT2RP2006004,		
	NT2RP2006042,	NT2RP2006269,	NT2RP1006512,	NT2RP2006580,	NT2RP3000169,	NT2RP3000171,
25	NT2RP3000304,	NT2RP3000436,	NT2RP3000460,	NT2RP3000616,	NT2RP3000676,	NT2RP3000721,
	NT2RP3000818,	NT2RP3000907,	NT2RP3000921,	NT2RP3001012,	NT2RP3001159,	NT2RP3001195,
	NT2RP3001240,	NT2RP3001271,	NT2RP3001322,	NT2RP3001388,	NT2RP3001560,	NT2RP3001592,
	NT2RP3001650,	NT2RP3001738,	NT2RP3001858,	NT2RP3002015,	NT2RP3002160,	NT2RP3002311,
	NT2RP3002342,	NT2RP3002411,	NT2RP3002737,	NT2RP3002790,	NT2RP3002836,	NT2RP3002900,
30	NT2RP3002958,	NT2RP3003000,	NT2RP3003076,	NT2RP3003354,	NT2RP3003532,	NT2RP3003535,
	NT2RP3003614,	NT2RP3004025,	NT2RP3004075,	NT2RP3004083,	NT2RP3004130,	NT2RP3004133,
	NT2RP3004309,	NT2RP3004345,	NT2RP3004406,	NT2RP3004481,	NT2RP3004552,	NT2RP3004625,
	NT2RP3004647,	NT2RP4001001,	NT2RP4001009,	NT2RP4001467,	NT2RP4001879,	NT2RP4002187,
	NT2RP4002451,	NT2RP4002750,	OVARC1000003,	OVARC1000105,	OVARC1000298,	OVARC1000307,
35	OVARC1000313,	OVARC1000410,	OVARC1000439,	OVARC1000553,	OVARC1000811,	OVARC1000873,
	OVARC1000956,	OVARC1001030,	OVARC1001163,	OVARC1001336,	OVARC1001570,	OVARC1001607,
	OVARC1001725,	OVARC1001991,	PLACE1000033,	PLACE1000231,	PLACE1000560,	PLACE1000740,
	PLACE1000912,	PLACE1000914,	PLACE1000927,	PLACE1001016,	PLACE1001123,	PLACE1001183,
	PLACE1001231,	PLACE1001340,	PLACE1001401,	PLACE1001407,	PLACE1001464,	PLACE1001516,
40	PLACE1001536,	PLACE1001564,	PLACE1001655,	PLACE1001795,		
	PLACE1001836,	PLACE1001918,	PLACE1001949,	PLACE1002080,	PLACE1002095,	PLACE1002355,
	PLACE1002374,	PLACE1002518,	PLACE1002547,	PLACE1002726,	PLACE1002905,	PLACE1002911,
	PLACE1002967,	PLACE1003407,	PLACE1003573,	PLACE1003737,	PLACE1003772,	PLACE1003839,
	PLACE1003845,	PLACE1003852,	PLACE1004279,	PLACE1004282,	PLACE1004441,	PLACE1004450,
45	PLACE1004482,	PLACE1004520,	PLACE1004630,	PLACE1004657,	PLACE1004648,	PLACE1004816,
	PLACE1005003,	PLACE1005005,	PLACE1005031,	PLACE1005383,	PLACE1005410,	PLACE1005426,
	PLACE1005544,	PLACE1005569,	PLACE1005660,	PLACE1005725,	PLACE1005745,	PLACE1005878,
	PLACE1005927,	PLACE1006071,	PLACE1006093,	PLACE1006208,	PLACE1006277,	PLACE1006290,
	PLACE1006443,	PLACE1006716,	PLACE1006809,	PLACE1006959,	PLACE1007081,	PLACE1007096,
50	PLACE1007296,	PLACE1007626,	PLACE1007845,	PLACE1007881,	PLACE1008359,	PLACE1008469,
	PLACE1008716,	PLACE1008744,	PLACE1008985,	PLACE1009067,	PLACE1009196,	PLACE1009279,
	PLACE1009527,	PLACE1009546,	PLACE1009600,	PLACE1009982,	PLACE1010011,	PLACE1010078,
	PLACE1010251,	PLACE1010445,	PLACE1010713,	PLACE1010784,	PLACE1010827,	PLACE1010968,
	PLACE1011116,	PLACE1011181,	PLACE1011236,	PLACE1011516,	PLACE1011708,	PLACE3000181,
55	PLACE3000213,	PLACE4000354,	SKNMC1000004,	SKNMC1000014,	SKNMC1000082,	THYRO1000036,
	THYRO1000099,	THYRO1000196,	THYRO1000400,	THYRO1000584,	THYRO1000678,	THYRO1000776,
	THYRO1000795,	THYRO1000956,	THYRO1001102,	THYRO1001113,		
	THYRO1001205,	THYRO1001237,	THYRO1001242,	THYRO1001266,	THYRO1001327,	THYRO1001456,

EP 1 130 094 A2

5 THYRO1001478, THYRO1001523, THYRO1001529, THYRO1001641, THYRO1001702, THYRO1001725,
 Y79AA1000207, Y79AA1000226, Y79AA1000270, Y79AA1000426, Y79AA1000521, Y79AA1000876,
 Y79AA1000888, Y79AA1000959, Y79AA1001013, Y79AA1001212, Y79AA1001264, Y79AA1001328,
 Y79AA1001426, Y79AA1001427, Y79AA1001430, Y79AA1001727, Y79AA1001787, Y79AA1001795,
 Y79AA1001799, Y79AA1001803, Y79AA1002022, Y79AA1002058, Y79AA1002129, Y79AA1002213,
 Y79AA1002373,

10 **[0036]** The following 146 clones were categorized into glycoprotein-associated proteins. The clones categorized into glycoprotein-associated proteins are those which matched the full-length sequences of Swiss-Prot database with the keyword "glycoprotein"; those which matched the data suggesting that the proteins are glycoprotein; or those which matched the full-length sequences of GenBank or UniGene database with similar description.

15 BNGH41000087, BNGH41000091, HEMBA1000349, HEMBA1000590, HEMBA1000745, HEMBA1000835,
 HEMBA1001221, HEMBA1001228, HEMBA1001621, HEMBA1002131, HEMBA1002178, HEMBA1002421,
 HEMBA1002767, HEMBA1003230, HEMBA1003392, HEMBA1004250, HEMBA1004391, HEMBA1004444,
 HEMBA1004505, HEMBA1005449, HEMBA1005522, HEMBA1005545, HEMBA1006707, HEMBA1006749,
 20 HEMBA1006902, HEMBB1000679, HEMBB1000881, HEMBB1001048, HEMBB1002120, HEMBB1002245,
 HEMBB1002427, MAMMA1000102, MAMMA1000591, MAMMA1000681, MAMMA1001043, MAMMA1001237,
 MAMMA1002070, MAMMA1002586, MAMMA1003126, NT2RM1000462, NT2RM1000580, NT2RM2001792,
 NT2RM2001818, NT2RM2001939, NT2RM2001941, NT2RM4000198, NT2RM4000284, NT2RM4000417,
 NT2RM4000648, NT2RM4000997, NT2RM4001325, NT2RM4002352, NT2RP1000613, NT2RP1000981,
 25 NT2RP1001004, NT2RP2000616, NT2RP2000694, NT2RP2000903, NT2RP2001480, NT2RP2001755,
 NT2RP2002533, NT2RP2003042, NT2RP2003210, NT2RP2004205, NT2RP2004606, NT2RP2005027,
 NT2RP2005181, NT2RP2005541, NT2RP2005597, NT2RP2005883, NT2RP2006004, NT2RP2006042,
 NT2RP2006269, NT2RP3000304, NT2RP3000616, NT2RP3000921, NT2RP3001650, NT2RP3002160,
 NT2RP3002737, NT2RP3002958, NT2RP3003000, NT2RP3003532, NT2RP3004130, NT2RP3004133,
 30 NT2RP3004481, NT2RP3004552, NT2RP3004640, NT2RP4000108, NT2RP4001467, NT2RP4002750,
 OVARC1000003, OVARC1000553, OVARC1000811, OVARC1000873, OVARC1001336, OVARC1001607,
 OVARC1001991, PLACE1000033, PLACE1000740, PLACE1001016,
 PLACE1001123, PLACE1001231, PLACE1001464, PLACE1001655, PLACE1001836, PLACE1002355,
 PLACE1002374, PLACE1002905, PLACE1002911, PLACE1003573, PLACE1003737, PLACE1003772,
 35 PLACE1003839, PLACE1004282, PLACE1004441, PLACE1004450, PLACE1004520, PLACE1004648,
 PLACE1005003, PLACE1005426, PLACE1006071, PLACE1006073, PLACE1006290, PLACE1007081,
 PLACE1007845, PLACE1008716, PLACE1008744, PLACE1008985, PLACE1010251, PLACE1010784,
 PLACE1010968, PLACE1011116, PLACE3000181, PLACE3000213, PLACE4000354, THYRO1000036,
 THYRO1000196, THYRO1000584, THYRO1000956, THYRO1001266, Y79AA1000270, Y79AA1000426,
 Y79AA1001727, Y79AA1001795, Y79AA1002022, Y79AA1002213,

40 **[0037]** The following 57 clones were categorized into signal transduction-associated proteins. The clones categorized into signal transduction-associated proteins are those which matched the full-length sequences of Swiss-Prot database with the keywords "serine/threonine-protein kinase", "tyrosine-protein kinase" or "SH3 domain"; those which matched the data suggesting that the proteins are signal transduction-associated proteins (for example, "ADP-ribosylation factor"); or those which matched the full-length sequences of GenBank or UniGene database with similar description; and, further, those which was similarly predicted to be signal transduction-associated proteins based on the matching data of Pfam.

45 HEMBA1000006, HEMBA1002195, HEMBA1002227, HEMBA1002551, HEMBA1005084, HEMBA1005929,
 HEMBA1006658, HEMBA1006916, MAMMA1000881, MAMMA1001150, MAMMA1001310, MAMMA1002142,
 NT2RM2001902, NT2RP1001020, NT2RP1001031, NT2RP2001469, NT2RP2001529, NT2RP2001769,
 NT2RP2003179, NT2RP2003545, NT2RP2004670, NT2RP3000011, NT2RP3000022, NT2RP3000172,
 NT2RP3000201, NT2RP3000820, NT2RP3003527, NT2RP3003849, NT2RP3003874, NT2RP3004067,
 NT2RP4000634, NT2RP4000962, OVARC1000255, OVARC1000529, OVARC1000916, OVARC1001338,
 50 OVARC1001569, PLACE1002329, PLACE1003135, PLACE1003598, PLACE1005519, PLACE1006208,
 PLACE1008282, PLACE1008297, PLACE1010081, PLACE1011364, PLACE1011824, THYRO1001457,
 THYRO1001593, THYRO1001700, THYRO1001770, Y79AA1000777, Y79AA1000967, Y79AA1002376,
 Y79AA1002381, HEMBB1000668, NT2RM4001377

55 **[0038]** The following 81 clones were categorized into transcription-associated proteins. The clones categorized into transcription-associated proteins are those which keywords "transcription regulation", "zinc finger" or "homeobox" matched the full-length sequences of Swiss-Prot database; those which matched the data suggesting that the proteins were transcription-associated proteins; or those which matched the full-length sequences of GenBank or UniGene database with similar description; and, further, those which was similarly predicted to be transcription-associated proteins based on the matching data of Pfam.

EP 1 130 094 A2

HEMBA1000462, HEMBA1000671, HEMBA1001297, HEMBA1001390, HEMBA1001886, HEMBA1002048,
HEMBA1003120, HEMBA1003497, HEMBA1004785, HEMBA1005230, HEMBA1005246, HEMBA1006276,
HEMBA1006572, HEMBA1007226, HEMBB1000106, HEMBB1000905, HEMBB1001959, HEMBB1002051,
HEMBB1002661, MAMMA1001094, MAMMA1001532, MAMMA1001615, NT2RM1000789, NT2RM2000632,
5 NT2RM2000773, NT2RM4000326, NT2RP1000271, NT2RP1000468, NT2RP2000092, NT2RP2000610,
NT2RP2000712, NT2RP2000739, NT2RP2001538, NT2RP2001662, NT2RP2001817, NT2RP2001948,
NT2RP2002564, NT2RP2002974, NT2RP2003138, NT2RP2003302, NT2RP2003940, NT2RP2004108,
NT2RP2004847, NT2RP2005247, NT2RP2005391, NT2RP2005535, NT2RP2005774, NT2RP2005941,
NT2RP2006092, NT2RP3000148, NT2RP3000232, NT2RP3000378, NT2RP3000652, NT2RP3001976,
10 NT2RP3004090, NT2RP3004119, NT2RP3004294, OVARC1001049, OVARC1001086, OVARC1001132,
OVARC1001807, PLACE1000258, PLACE1000442, PLACE1000907, PLACE1003529, PLACE1004166,
PLACE1004168, PLACE1004887, PLACE1005250, PLACE1005682, PLACE1006079, PLACE1008549,
PLACE1011407, PLACE1011978, THYRO1000580, Y79AA1000030, Y79AA1001090, Y79AA1001523,
Y79AA1002334, Y79AA1002378, HEMBB1002302,

15 **[0039]** The following 85 clones were categorized into disease-associated proteins. The clones categorized into disease-associated proteins are those which matched the full-length sequences of Swiss-Prot database with the keywords "disease mutation" or "syndrome"; those which matched the data suggesting that the proteins are disease-associated proteins; or those which matched the full-length sequences of Swiss-Prot database and GenBank or UniGene database where the matched sequences are those of genes or proteins which had been deposited in the database of Online
20 Mendelian Inheritance in Man (OMIM) (<http://www.ncbi.nlm.nih.gov/Omim/>), which is a database of human genes and diseases.

BNGH41000020, HEMBA1000349, HEMBA1000590, HEMBA1000671, HEMBA1000835, HEMBA1001184,
HEMBA1001228, HEMBA1001886, HEMBA1003120, HEMBA1004250, HEMBA1005246, HEMBA1005267,
HEMBA1006707, HEMBA1006749, HEMBA1006902, HEMBA1006916, HEMBA1007013, HEMBB1002120,
25 MAMMA1000204, MAMMA1002080, NT2RM2000632, NT2RM2001126, NT2RM2001558, NT2RP1000271,
NT2RP1000465, NT2RP1000579, NT2RP2000447, NT2RP2000514, NT2RP2000739, NT2RP2001223,
NT2RP2001529, NT2RP2001562, NT2RP2002674, NT2RP2003369, NT2RP2004108, NT2RP2004205,
NT2RP2005535, NT2RP2005941, NT2RP2006004, NT2RP3000059, NT2RP3000125, NT2RP3000201,
NT2RP3000232, NT2RP3000616, NT2RP3000677, NT2RP3000838, NT2RP3000921, NT2RP3001542,
30 NT2RP3002286, NT2RP3002721, NT2RP3002737, NT2RP3002738, NT2RP3004481, OVARC1000208,
OVARC1000275, OVARC1000331, OVARC1000410, OVARC1001086, OVARC1001132, OVARC1001607,
OVARC1001725, OVARC1001952, PLACE1000258, PLACE1000442, PLACE1000907, PLACE1001100,
PLACE1001500, PLACE1002905, PLACE1002967, PLACE1003407, PLACE1003428, PLACE1005005,
PLACE1005239, PLACE1005815, PLACE1007028, PLACE1008716, PLACE1011407, PLACE1011978,
35 PLACE2000118, THYPO1000580, THYRO1000866, THYRO1001071, THYRO1001478, Y79AA1001062,
Y79AA1001530,

[0040] It is unclear, by the analyses so far, whether or not the remaining 212 clones encode proteins belonging to any of the categories of secretory and/or membrane proteins, glycoprotein-associated proteins, signal transduction-associated proteins, transcription-associated proteins or disease-associated proteins. Nonetheless, it is still possible
40 for these clones to encode secretory and/or membrane proteins, glycoprotein-associated proteins, signal transduction-associated proteins, transcription-associated proteins, or disease-associated proteins. On the other hand, some of these clones can be presumed to have functions other than those as secretory and/or membrane proteins, glycoprotein-associated proteins, signal transduction-associated proteins, transcription-associated proteins and disease-associated proteins.

45 **[0041]** Among the 212 clones, the following clones presumably belong to the categories of enzymes and/or metabolism-associated proteins, cell division- and/or cell proliferation-associated proteins, cytoskeleton-associated proteins, nuclear proteins, DNA- and/or RNA-binding proteins, ATP- and/or GTP-binding proteins, protein synthesis- and/or protein transport-associated proteins, or cellular defense-associated proteins, although it is unclear whether or not the clones belong to any of the categories of secretory and/or transmembrane proteins, glycoprotein-associated proteins,
50 signal transduction-associated proteins, transcription-associated proteins, and disease-associated proteins.

[0042] The following 10 clones presumably belong to the category of enzymes and/or metabolism-associated proteins. The clones herein defined as clones presumably belonging to the category of enzymes and/or metabolism-associated proteins matched data containing keywords such as "metabolism", "oxidoreductase" and "E.C. No. (Enzyme commission number)".

55 HEMBA1003315, HEMBB1002465, MAMMA1000614, NT2RP2000178, NT2RP2001388, NT2RP2001903, NT2RP2002304, NT2RP2005878, NT2RP3001685, PLACE1006219

[0043] The following 4 clones presumably belong to the category of cell division- and/or cell proliferation-associated proteins. The cDNA clones were herein defined as clones presumably belonging to the category of cell division- and/

EP 1 130 094 A2

or cell proliferation-associated proteins matched data containing keywords such as "cell division", "cell cycle", "mitosis", "chromosomal protein", "cell growth" and "apoptosis".

MAMMA1000403, NT2RM2000497, NT2RP2000394, Y79AA1002121

[0044] The following 6 clones presumably belong to the category of cytoskeleton-associated proteins. The cDNA clones were herein defined as clones presumably belonging to the category of cytoskeleton-associated proteins matched data containing keywords such as "structural protein", "cytoskeleton", "actin-binding" and "microtubules".

MAMMA 1001609, NT2RM2000589, NT2RP3000063, PLACE 1004078, PLACE 1004492, PLACE 1008657

[0045] The following 7 clones presumably belong to the category of nuclear proteins. The cDNA clones were herein defined as clones presumably belonging to the category of nuclear proteins matched data containing keywords such as "nuclear protein".

HEMBA1001878, HEMBA1002992, MAMMA1000614, NT2RM4000965, NT2RM2001738, NT2RP2001388, Y79AA1002121

[0046] The following 5 clones presumably belong to the category of DNA- and/or RNA-binding proteins. The cDNA clones were herein defined as clones presumably belonging to the category of DNA- and/or RNA-binding proteins matched data containing keywords such as "DNA-binding" and "RNA-binding".

HEMBA1003072, HEMBA1006770, HEMBA1007332, NT2RM2000497, Y79AA1002121

[0047] The following 7 clones presumably belong to the category of ATP- and/or GTP-binding proteins. The cDNA clones were herein defined as clones presumably belonging to the category of ATP- and/or GTP-binding proteins matched data containing keywords such as "ATP-binding" and "GTP-binding".

HEMBA1002316, MAMMA1001609, NT2RM2000306, NT2RM2000497, NT2RP2000178, NT2RP3003729, PLACE1004305

[0048] The following 7 clones presumably belong to the category of protein synthesis- and/or protein transport-associated proteins. The cDNA clones were herein defined as clones presumably belonging to the category of protein synthesis-associated and/or protein transport-associated proteins matched data containing keywords such as "translation regulation", "protein biosynthesis", "amino-acid biosynthesis", "ribosomal protein", "protein transport" and "signal recognition particle".

NT2RM4000965, NT2RP2005069, NT2RP3000481, NT2RP3000789, NT2RP4001877, OVARC1001833, OVARC1002058,

[0049] The following 1 clone presumably belongs to the category of cellular defense-associated proteins. The cDNA clones were herein defined as clones presumably belonging to the category of cellular defense-associated proteins matched data containing keywords such as "heat shock", "DNA repair" and "DNA damage".

PLACE1005539

[0050] Although it is unclear whether or not 26 out of 174 clones other than the above-mentioned clones belong to any of the above-described categories, these clones are predicted to have some functions, based on the homology search using their full-length sequences thereof. The clone names and the gene definitions found in the result of homology search are shown below, separated with a double-slash mark, //.

HEMBA1000634//Homo sapiens T-cell activation protein (PGR1) gene, complete cds.

HEMBA1002524//Human MHC Class I region proline rich protein mRNA, complete cds.

HEMBA1003399//MVP1 PROTEIN.

HEMBA1005489//Mus musculus semaphorin cytoplasmic domain-associated protein 3A (Semcap3) mRNA, complete cds.

HEMBB1000542//Mus musculus bromodomain-containing protein BP75 mRNA, complete cds.

MAMMA1000788//Bos taurus P14 (p14) mRNA, complete cds.

MAMMA1002128//ABC1 PROTEIN HOMOLOG PRECURSOR.

NT2RM2000514//Homo Sapiens F-box protein Fbx21 (FBX21) mRNA, complete cds.

NT2RM2000622//Mus musculus F-box protein FBL10 mRNA, partial cds.

NT2RM4000100//Homo sapiens Leman coiled-coil protein (LCCP) mRNA, complete cds.

NT2RP2005425//Homo sapiens mRNA for AKAP450 protein.

NT2RP3001170//Mus musculus activity-dependent neuroprotective protein (Adnp) mRNA, complete cds.

NT2RP3002571//Bos taurus mRNA for lyncein.

NT2RP3004557//Human Ki nuclear autoantigen mRNA, complete cds.

OVARC1001596//Homo sapiens Arf-like 2 binding protein BART1 mRNA, complete cds.

PLACE1002153//Homo sapiens TACC2 protein (TACC2) mRNA, partial cds.

PLACE1003163//Homo sapiens DBI-related protein mRNA, complete cds.

PLACE1005736//Human mRNA for BAS-GRIP protein.

PLACE1007702//Mus musculus TRA1 mRNA, complete cds.

PLACE1011045//Homo sapiens E1-like protein mRNA, complete cds.

EP 1 130 094 A2

THYRO1000061//Mus musculus mRNA for UBE-1c1, UBE-1c2, UBE-1c3, complete cds.
 THYRO1000964//Drosophila melanogaster Pelle associated protein Pellino (Pli) mRNA, complete cds.
 Y79AA1000776//Mus musculus mRNA for GSG1, complete cds.
 Y79AA1001056//Homo sapiens MAID protein mRNA, complete cds.
 5 Y79AA1001272//Homo sapiens retinoic acid repressible protein (RARG-1) mRNA, complete cds.
 Y79AA1001793//Mus musculus mRNA for GSG1, complete cds.

[0051] So far, useful information for presuming the functions are unavailable for the remaining 148 clones, whose names are listed below.

10 HEMBA1000275, HEMBA1000300, HEMBA1000443, HEMBA1000875, HEMBA1000907, HEMBA1001272,
 HEMBA1001296, HEMBA1001563, HEMBA1002164, HEMBA1002239, HEMBA1002985, HEMBA1003294,
 HEMBA1003487, HEMBA1004007, HEMBA1004067, HEMBA1004085, HEMBA1004952, HEMBA1004971,
 HEMBA1005145, HEMBA1005430, HEMBA1005913, HEMBA1006016, HEMBA1006517, HEMBA1006544,
 HEMBA1006912, HEMBA1007057, HEMBA1007063, HEMBA1007291, HEMBB1000276, HEMBB1000309,
 15 HEMBB1000642, HEMBB1001200, HEMBB1001547, HEMBB1002039, HEMBB1002228, HEMBB1002663,
 MAMMA1000046, MAMMA1000449, MAMMA1000528, MAMMA1000652, MAMMA1000706, MAMMA1000814,
 MAMMA1001066, MAMMA1001284, MAMMA1001623, MAMMA1001634, MAMMA1001901, MAMMA1002087,
 MAMMA1002205, MAMMA1002224, NT2RM2000582, NT2RM2001643, NT2RM4000115, NT2RM4000295,
 NT2RM4001321, NT2RP1000002, NT2RP1000239, NT2RP1000679, NT2RP1000740, NT2RP1000903,
 20 NT2RP2000240, NT2RP2001878, NT2RP2001921, NT2RP2002015, NT2RP2002409, NT2RP2002510,
 NT2RP2003599, NT2RP2003931, NT2RP2004069, NT2RP2004141, NT2RP2004447, NT2RP2004837,
 NT2RP2005514, NT2RP2005632, NT2RP2005887, NT2RP2006099, NT2RP2006134, NT2RP3000427,
 NT2RP3000444, NT2RP3000645, NT2RP3000871, NT2RP3001044, NT2RP3001061, NT2RP3001754,
 NT2RP3002281, NT2RP3002324, NT2RP3002353, NT2RP3002409, NT2RP3002448, NT2RP3002664,
 25 NT2RP3002887, NT2RP3002983, NT2RP3003448, NT2RP3003469, NT2RP3003473, NT2RP3003559,
 NT2RP3003963, NT2RP3004000, NT2RP3004202, NT2RP3004321,
 NT2RP3004355, NT2RP3004374, NT2RP4002715, OVARC1000090, OVARC1000137, OVARC1000467,
 OVARC1000775, OVARC1000853, OVARC1000995, OVARC1001222, OVARC1001260, OVARC1001727,
 OVARC1002178, PLACE1000986, PLACE1001114, PLACE1001229, PLACE1001788, PLACE1003438,
 30 PLACE1003460, PLACE1003644, PLACE1004028, PLACE1004199, PLACE1004519, PLACE1005601,
 PLACE1005669, PLACE1005768, PLACE1006515, PLACE1006786, PLACE1007040, PLACE1007077,
 PLACE1007591, PLACE1007971, PLACE1008984, PLACE1009735, PLACE2000219, PLACE4000455,
 THYRO1000846, THYRO1000999, THYRO1001063, THYRO1001128, THYRO1001471, THYRO1001495,
 THYRO1001608, THYRO1001803, Y79AA1000127, Y79AA1000750, Y79AA1001592, Y79AA1001863,

35 **[0052]** In the 437 clones categorized into secretory and/or membrane proteins by using their full-length sequences, 410 clones were also predicted to encode proteins having functions of secretory and/or membrane proteins by using their partial nucleotide sequences. In the 146 clones categorized into glycoprotein-associated proteins by using their full-length sequences, 124 clones were also predicted to encode proteins having functions of glycoprotein-associated proteins by using their partial nucleotide sequences. In the 57 clones categorized into signal transduction-associated proteins by using their full-length sequences, 46 clones were also predicted to encode proteins having functions of signal transduction-associated proteins by using their partial nucleotide sequences. In the 81 clones categorized into transcription-associated proteins by using their full-length sequences, 57 clones were also predicted to encode proteins having functions of transcription-associated proteins by using their partial nucleotide sequences. In the 85 clones categorized into disease-associated proteins by using their full-length sequences, 6 clones were also predicted to encode proteins having functions of disease-associated proteins by using their partial nucleotide sequences. The number of clones, which were predicted to encode disease-associated proteins based on the full-length nucleotide sequences, is much greater than that predicted based on the partial sequences. The reason is that the full-length sequences were categorized by using the data found in the OMIM database into the category of disease-associated proteins.

40 **[0053]** In some cases, the predicted functions based on the partial sequences are different from those based on the full-length sequences. The reason is that a protein does not always belong solely to a single category of the above-described functional categories, and therefore, it is possible for the protein to belong to both of the predicted functional categories. Besides, additional functions can be found for the clones classified into these functional categories by further analyses.

50 **[0054]** The following list shows the cDNA clones predicted and selected on the basis of the partial sequences (5' sequences) as cDNAs encoding secretory and/or membrane proteins, glycoprotein-associated proteins, signal transduction-associated proteins, transcription-associated proteins, or disease-associated proteins.

[0055] The clones that are selected by the score in the ATGpr and by the PSORT for the existence of a signal sequence can be expected to encode a secretory or membrane protein since they are predicted to possess the secretion

EP 1 130 094 A2

signal or a transmembrane region. The clones that are selected by the score in the ATGpr and by the PSORT for the existence of a signal sequence are listed below (254 clones).

5 HEMBA1000300 HEMBA1000713 HEMBA1000907
 HEMBA1000962 HEMBA1001272 HEMBA1001297
 HEMBA1002164 HEMBA1002239 HEMBA1002420
 HEMBA1002421 HEMBA1003101 HEMBA1003294
 HEMBA1003399 HEMBA1003602 HEMBA1003732
 HEMBA1004110 HEMBA1004797 HEMBA1005430
 10 HEMBA1006016 HEMBA1006171 HEMBA1006311
 HEMBA1006335 HEMBA1006357 HEMBA1006572
 HEMBA1006658 HEMBA1006707 HEMBA1006902
 HEMBA1006960 HEMBA1007013 HEMBB1000276
 HEMBB1000447 HEMBB1000567 HEMBB1000642
 15 HEMBB1000905 HEMBB1001200 HEMBB1001407
 HEMBB1001530 HEMBB1001547 HEMBB1001978
 HEMBB1002162 HEMBB1002228 HEMBB1002245
 HEMBB1002427 HEMBB1002465 HEMBB1002663
 HEMBB1002693 MAMMA1000046 MAMMA1000102
 20 MAMMA1000118 MAMMA1000141 MAMMA1000449
 MAMMA1000457 MAMMA1000591 MAMMA1000652
 MAMMA1000681 MAMMA1000986 MAMMA1000994
 MAMMA1001043 MAMMA1001141 MAMMA1001284
 MAMMA1001310 MAMMA1001344 MAMMA1001893
 25 MAMMA1001901 MAMMA1001957 MAMMA1002070
 MAMMA1002087 MAMMA1002165 MAMMA1002205
 MAMMA1002224 MAMMA1002633 NT2RM2000241
 NT2RM2000306 NT2RM2000410 NT2RM2000514
 NT2RM2001643 NT2RM2001941 NT2RM4000115
 30 NT2RM4000997 NT2RM4001321 NT2RM4001325
 NT2RM4001768 NT2RP1000050 NT2RP1000448
 NT2RP1000903 NT2RP1001563 NT2RP2000479
 NT2RP2001495 NT2RP2001915 NT2RP2001948
 NT2RP2002015 NT2RP2002063 NT2RP2002304
 35 NT2RP2002674 NT2RP2002721 NT2RP2003383
 NT2RP2003469 NT2RP2003593 NT2RP2003599
 NT2RP2003655 NT2RP2003664 NT2RP2004179
 NT2RP2004447 NT2RP2004495 NT2RP2004524
 NT2RP2004556 NT2RP2004837 NT2RP2005027
 40 NT2RP2005463 NT2RP2005514 NT2RP2005887
 NT2RP2006042 NT2RP2006269 NT2RP3000169
 NT2RP3000460 NT2RP3000481 NT2RP3000645
 NT2RP3000789 NT2RP3000818 NT2RP3001012
 NT2RP3001044 NT2RP3001195 NT2RP3001560
 45 NT2RP3001685 NT2RP3001858 NT2RP3002160
 NT2RP3002281 NT2RP3002721 NT2RP3002836
 NT2RP3002958 NT2RP3003076 NT2RP3003354
 NT2RP3003469 NT2RP3003535 NT2RP3003559
 NT2RP3003963 NT2RP3004000 NT2RP3004083
 50 NT2RP3004133 NT2RP3004309 NT2RP3004321
 NT2RP3004355 NT2RP3004374 NT2RP4001001
 NT2RP4001879 NT2RP4002451 NT2RP4002715
 OVARC1000208 OVARC1000298 OVARC1000439
 OVARC1000775 OVARC1000811 OVARC1000853
 55 OVARC1001222 OVARC1001727 OVARC1001807
 OVARC1001833 PLACE1000231 PLACE1000560
 PLACE1000740 PLACE1000912 PLACE1000914
 PLACE1000927 PLACE1000986 PLACE1001100

EP 1 130 094 A2

PLACE1001183 PLACE1001229 PLACE1001407
 PLACE1001536 PLACE1001788 PLACE1002080
 PLACE1002095 PLACE1002374 PLACE1002518
 PLACE1003407 PLACE1003428 PLACE1003460
 5 PLACE1003839 PLACE1003845 PLACE1004028
 PLACE1004199 PLACE1004282 PLACE1004305
 PLACE1004482 PLACE1004637 PLACE1005005
 PLACE1005250 PLACE1005383 PLACE1005410
 PLACE1005544 PLACE1005569 PLACE1005601
 10 PLACE1005660 PLACE1005669 PLACE1005725
 PLACE1005768 PLACE1005927 PLACE1006079
 PLACE1006093 PLACE1006219 PLACE1006277
 PLACE1006443 PLACE1006786 PLACE1006809
 PLACE1007040 PLACE1007096 PLACE1007296
 15 PLACE1007626 PLACE1007971 PLACE1008469
 PLACE1008984 PLACE1008985 PLACE1009067
 PLACE1009196 PLACE1009527 PLACE1009982
 PLACE1010251 PLACE1011236 PLACE2000219
 PLACE4000455 SKNMC1000004 SKNMC1000014
 20 THYRO1000036 THYRO1000099 THYRO1000196
 THYRO1000795 THYRO1000999 THYRO1001237
 THYRO1001327 THYRO1001478 THYRO1001495
 THYRO1001523 THYRO1001702 THYRO1001725
 Y79AA1000226 Y79AA1000270 Y79AA1000426
 25 Y79AA1000521 Y79AA1000776 Y79AA1000959
 Y79AA1001013 Y79AA1001056 Y79AA1001264
 Y79AA1001328 Y79AA1001427 Y79AA1001430
 Y79AA1001530 Y79AA1001592 Y79AA1001793
 Y79AA1001795 Y79M1001803 Y79AA1001863
 30 Y79AA1002022 Y79AA1002373

[0056] In the example mentioned below, the 254 clones as described above were categorized into three groups according to their maximal value in the ATGpr and the result in the PSORT, which are shown in Table 7-10, 11, 12 (246 clones), and Table 13, 14, 15 (8 clones). In the tables, the name of clone, indicate the name of the clone that was selected by the ATGpr and the PSORT; the name of sequence indicates the name of the 5'-end sequence of the clone on the left; the maximal ATGpr score indicates the maximal ATGpr1 score of the 5'-end sequence shown on the left; and signal indicates the presence of the signal sequence according to the prediction by the PSORT. In addition, the representative sequence is the sequence that has the longest sequence among the cluster in which the 5'-end sequence on the left was included. The maximal ATGpr score and signal on the right indicate the maximal ATGpr1 score of the representative sequence, and the presence of a signal sequence in the representative sequence according to the prediction by the PSORT, respectively. The 170 clones shown in Table 7-10, having the maximal score in the ATGpr1 higher than 0.5, and predicted to possess a signal sequence by the PSORT, are very likely to be full-length and encode a secretory or membrane protein. The 35 clones in Table 11, which have the maximal score in the ATGpr1 0.3 or higher and less than 0.5, and predicted to have a signal sequence, are also as well. And, the 41 clones in Table 12, having the maximal score in the ATGpr1 0 or higher and less than 0.3, and predicted to have a signal sequence, are likely to be full-length and encode a secretory or membrane protein.

[0057] The 8 clones in Table 13 (4 clones), Table 14 (2 clones), and Table 15 (2 clones) have the maximal score in the ATGpr1 0.5 or higher, 0.3 or higher and less than 0.5, and 0 or higher and less than 0.3, respectively, and are predicted to have no signal sequence by the PSORT. However, these clones contain a region that is recognized by the PSORT to be a signal sequence within the representative sequence composing the same cluster. Thus, the clones were judged as a full-length clone which encodes a membrane protein, especially.

[0058] The clones selected by the score in the ATGpr and by the keywords in the top hit data in the SwissProt are likely to encode a secretory or membrane protein, or proteins with functions associated to signal transduction, glyco-protein, transcription, and diseases according to the respective keywords. These 659 clones are shown below. Here, top hit data is defined to be data of known amino acid sequence which is identified to be the most homologous sequence in homology search using the SwissProt.

BNGH41000020 BNGH41000087 BNGH41000091

EP 1 130 094 A2

HEMBA1000006 HEMBA1000121 HEMBN1000128
 HEMBA1000275 HEMBA1000349 HEMBA1000443
 HEMBA1000462 HEMBA1000477 HEMBA1000590
 HEMBA1000634 HEMBA1000671 HEMBA1000732
 5 HEMBA1000745 HEMBA1000835 HEMBA1000875
 HEMBA1000907 HEMBA1000940 HEMBA1001184
 HEMBA1001221 HEMBA1001228 HEMBA1001296
 HEMBA1001390 HEMBA1001563 HEMBA1001621
 HEMBA1001878 HEMBA1001886 HEMBA1002048
 10 HEMBA1002131 HEMBA1002163 HEMBA1002164
 HEMBA1002167 HEMBA1002178 HEMBA1002195
 HEMBA1002227 HEMBA1002316 HEMBA1002421
 HEMBA1002524 HEMBA1002551 HEMBA1002767
 HEMBA1002985 HEMBA1002992 HEMBA1003047
 15 HEMBA1003072 HEMBA1003101 HEMBA1003120
 HEMBA1003230 HEMBA1003315 HEMBA1003392
 HEMBA1003487 HEMBA1003497 HEMBA1003530
 HEMBA1003945 HEMBA1004007 HEMBA1004067
 HEMBA1004085 HEMBA1004250 HEMBA1004391
 20 HEMBA1004444 HEMBA1004454 HEMBA1004505
 HEMBA1004785 HEMBA1004797 HEMBA1004952
 HEMBA1004971 HEMBA1004982 HEMBA1005070
 HEMBA1005084 HEMBA1005145 HEMBA1005230
 HEMBA1005246 HEMBA1005267 HEMBA1005337
 25 HEMBA1005449 HEMBA1005489 HEMBA1005522
 HEMBA1005545 HEMBA1005698 HEMBA1005913
 HEMBA1005929 HEMBA1005945 HEMBA1006276
 HEMBA1006299 HEMBA1006335 HEMBA1006430
 HEMBA1006482 HEMBA1006517 HEMBA1006544
 30 HEMBA1006572 HEMBA1006707 HEMBA1006724
 HEMBA1006749 HEMBA1006770 HEMBA1006902
 HEMBA1006912 HEMBA1006916 HEMBA1007013
 HEMBA1007057 HEMBA1007063 HEMBA1007226
 HEMBA1007241 HEMBA1007291 HEMBA1007332
 35 HEMBB1000106 HEMBB1000309 HEMBB1000407
 HEMBB1000447 HEMBB1000542 HEMBB1000567
 HEMBB1000668 HEMBB1000679 HEMBB1000881
 HEMBB1001026 HEMBB1001048 HEMBB1001200
 HEMBB1001573 HEMBB1001847 HEMBB1001959
 40 HEMBB1002039 HEMBB1002041 HEMBB1002051
 HEMBB1002120 HEMBB1002302 HEMBB1002427
 HEMBB1002661 MAMMA1000106 MAMMA1000204
 MAMMA1000226 MAMMA1000403 MAMMA1000473
 MAMMA1000496 MAMMA1000528 MAMMA1000591
 45 MAMMA1000614 MAMMA1000681 MAMMA1000706
 MAMMA1000788 MAMMA1000810 MAMMA1000814
 MAMMA1000881 MAMMA1001043 MAMMA1001066
 MAMMA1001094 MAMMA1001150 MAMMA1001237
 MAMMA1001418 MAMMA1001532 MAMMA1001609
 50 MAMMA1001615 MAMMA1001623 MAMMA1001634
 MAMMA1001893 MAMMA1001957 MAMMA1001978
 MAMMA1002070 MAMMA1002080 MAMMA1002091
 MAMMA1002095 MAMMA1002128 MAMMA1002142
 MAMMA1002165 MAMMA1002234 MAMMA1002586
 55 MAMMA1002633 MAMMA1003126 NT2RM1000407
 NT2RM1000462 NT2RM1000542 NT2RM1000580
 NT2RM1000789 NT2RM1000855 NT2RM1000858
 NT2RM1000899 NT2RM2000410 NT2RM2000423

EP 1 130 094 A2

NT2RM2000497 NT2RM2000565 NT2RM2000582
 NT2RM2000589 NT2RM2000622 NT2RM2000632
 NT2RM2000773 NT2RM2001126 NT2RM2001558
 NT2RM2001626 NT2RM2001738 NT2RM2001767
 5 NT2RM2001792 NT2RM2001818 NT2RM2001902
 NT2RM2001939 NT2RM2001941 NT2RM4000100
 NT2RM4000198 NT2RM4000284 NT2RM4000295
 NT2RM4000326 NT2RM4000417 NT2RM4000444
 NT2RM4000587 NT2RM4000593 NT2RM4000648
 10 NT2RM4000761 NT2RM4000965 NT2RM4001377
 NT2RM4001735 NT2RM4001843 NT2RM4002352
 NT2RP1000002 NT2RP1000050 NT2RP1000181
 NT2RP1000239 NT2RP1000261 NT2RP1000271
 NT2RP1000300 NT2RP1000325 NT2RP1000465
 15 NT2RP1000468 NT2RP1000551 NT2RP1000579
 NT2RP1000613 NT2RP1000679 NT2RP1000740
 NT2RP1000981 NT2RP1001004 NT2RP1001020
 NT2RP1001031 NT2RP2000092 NT2RP2000178
 NT2RP2000240 NT2RP2000394 NT2RP2000447
 20 NT2RP2000514 NT2RP2000533 NT2RP2000610
 NT2RP2000616 NT2RP2000649 NT2RP2000663
 NT2RP2000694 NT2RP2000712 NT2RP2000739
 NT2RP2000818 NT2RP2000903 NT2RP2001200
 NT2RP2001223 NT2RP2001276 NT2RP2001388
 25 NT2RP2001469 NT2RP2001480 NT2RP2001495
 NT2RP2001514 NT2RP2001529 NT2RP2001538
 NT2RP2001562 NT2RP2001662 NT2RP2001755
 NT2RP2001769 NT2RP2001817 NT2RP2001878
 NT2RP2001903 NT2RP2001921 NT2RP2001948
 30 NT2RP2001956 NT2RP2002063 NT2RP2002188
 NT2RP2002232 NT2RP2002304 NT2RP2002409
 NT2RP2002510 NT2RP2002527 NT2RP2002533
 NT2RP2002564 NT2RP2002824 NT2RP2002942
 NT2RP2002974 NT2RP2002976 NT2RP2003042
 35 NT2RP2003138 NT2RP2003179 NT2RP2003210
 NT2RP2003302 NT2RP2003369 NT2RP2003390
 NT2RP2003469 NT2RP2003545 NT2RP2003593
 NT2RP2003655 NT2RP2003664 NT2RP2003931
 NT2RP2003940 NT2RP2003950 NT2RP2004069
 40 NT2RP2004108 NT2RP2004141 NT2RP2004205
 NT2RP2004447 NT2RP2004606 NT2RP2004648
 NT2RP2004670 NT2RP2004794 NT2RP2004847
 NT2RP2005069 NT2RP2005163 NT2RP2005181
 NT2RP2005247 NT2RP2005378 NT2RP2005391
 45 NT2RP2005425 NT2RP2005535 NT2RP2005541
 NT2RP2005597 NT2RP2005632 NT2RP2005666
 NT2RP2005774 NT2RP2005878 NT2RP2005883
 NT2RP2005941 NT2RP2005994 NT2RP2006004
 NT2RP2006042 NT2RP2006092 NT2RP2006099
 50 NT2RP2006134 NT2RP2006269 NT2RP2006512
 NT2RP3000011 NT2RP3000022 NT2RP3000059
 NT2RP3000063 NT2RP3000125 NT2RP3000148
 NT2RP3000171 NT2RP3000172 NT2RP3000201
 NT2RP3000232 NT2RP3000304 NT2RP3000378
 55 NT2RP3000427 NT2RP3000436 NT2RP3000444
 NT2RP3000481 NT2RP3000616 NT2RP3000645
 NT2RP3000652 NT2RP3000676 NT2RP3000677
 NT2RP3000721 NT2RP3000820 NT2RP3000838

EP 1 130 094 A2

NT2RP3000871 NT2RP3000907 NT2RP3000921
 NT2RP3001012 NT2RP3001061 NT2RP3001159
 NT2RP3001170 NT2RP3001195 NT2RP3001240
 NT2RP3001271 NT2RP3001322 NT2RP3001388
 5 NT2RP3001542 NT2RP3001560 NT2RP3001592
 NT2RP3001650 NT2RP3001738 NT2RP3001754
 NT2RP3001976 NT2RP3002015 NT2RP3002160
 NT2RP3002286 NT2RP3002311 NT2RP3002324
 NT2RP3002342 NT2RP3002353 NT2RP3002409
 10 NT2RP3002411 NT2RP3002448 NT2RP3002571
 NT2RP3002664 NT2RP3002737 NT2RP3002738
 NT2RP3002790 NT2RP3002836 NT2RP3002887
 NT2RP3002900 NT2RP3002958 NT2RP3002983
 NT2RP3003000 NT2RP3003076 NT2RP3003354
 15 NT2RP3003448 NT2RP3003473 NT2RP3003527
 NT2RP3003532 NT2RP3003614 NT2RP3003729
 NT2RP3003849 NT2RP3003874 NT2RP3003939
 NT2RP3004025 NT2RP3004067 NT2RP3004075
 NT2RP3004090 NT2RP3004119 NT2RP3004130
 20 NT2RP3004133 NT2RP3004202 NT2RP3004294
 NT2RP3004309 NT2RP3004345 NT2RP3004406
 NT2RP3004481 NT2RP3004552 NT2RP3004557
 NT2RP3004625 NT2RP3004640 NT2RP3004647
 NT2RP4000108 NT2RP4000634 NT2RP4000962
 25 NT2RP4001009 NT2RP4001467 NT2RP4001877
 NT2RP4001879 NT2RP4002187 NT2RP4002451
 NT2RP4002750 OVARC1000003 OVARC1000090
 OVARC1000105 OVARC1000137 OVARC1000255
 OVARC1000275 OVARC1000307 OVARC1000313
 30 OVARC1000331 OVARC1000410 OVARC1000439
 OVARC1000467 OVARC1000529 OVARC1000553
 OVARC1000873 OVARC1000916 OVARC1000956
 OVARC1000995 OVARC1001030 OVARC1001049
 OVARC1001086 OVARC1001132 OVARC1001163
 35 OVARC1001222 OVARC1001260 OVARC1001336
 OVARC1001338 OVARC1001569 OVARC1001570
 OVARC1001596 OVARC1001607 OVARC1001725
 OVARC1001952 OVARC1001991 OVARC1002058
 OVARC1002178 PLACE1000033 PLACE1000258
 40 PLACE1000442 PLACE1000740 PLACE1000907
 PLACE1001016 PLACE1001114 PLACE1001123
 PLACE1001231 PLACE1001340 PLACE1001401
 PLACE1001407 PLACE1001464 PLACE1001500
 PLACE1001516 PLACE1001564 PLACE1001655
 45 PLACE1001795 PLACE1001836 PLACE1001918
 PLACE1001949 PLACE1002080 PLACE1002095
 PLACE1002153 PLACE1002329 PLACE1002355
 PLACE1002374 PLACE1002547 PLACE1002726
 PLACE1002905 PLACE1002911 PLACE1002967
 50 PLACE1003135 PLACE1003163 PLACE1003428
 PLACE1003438 PLACE1003460 PLACE1003529
 PLACE1003573 PLACE1003598 PLACE1003644
 PLACE1003737 PLACE1003772 PLACE1003852
 PLACE1004078 PLACE1004166 PLACE1004168
 55 PLACE1004279 PLACE1004441 PLACE1004450
 PLACE1004482 PLACE1004492 PLACE1004519
 PLACE1004520 PLACE1004630 PLACE1004648
 PLACE1004816 PLACE1004887 PLACE1005003

EP 1 130 094 A2

PLACE1005031 PLACE1005239 PLACE1005383
 PLACE1005426 PLACE1005519 PLACE1005539
 PLACE1005544 PLACE1005569 PLACE1005682
 PLACE1005736 PLACE1005745 PLACE1005815
 5 PLACE1005878 PLACE1005927 PLACE1006071
 PLACE1006073 PLACE1006208 PLACE1006277
 PLACE1006290 PLACE1006443 PLACE1006515
 PLACE1006716 PLACE1006959 PLACE1007028
 PLACE1007077 PLACE1007081 PLACE1007096
 10 PLACE1007296 PLACE1007591 PLACE1007702
 PLACE1007845 PLACE1007881 PLACE1008282
 PLACE1008297 PLACE1008359 PLACE1008469
 PLACE1008549 PLACE1008657 PLACE1008716
 PLACE1008744 PLACE1008984 PLACE1008985
 15 PLACE1009279 PLACE1009527 PLACE1009546
 PLACE1009600 PLACE1009735 PLACE1010011
 PLACE1010078 PLACE1010081 PLACE1010251
 PLACE1010445 PLACE1010713 PLACE1010784
 PLACE1010827 PLACE1010968 PLACE1011045
 20 PLACE1011116 PLACE1011181 PLACE1011236
 PLACE1011364 PLACE1011407 PLACE1011516
 PLACE1011708 PLACE1011824 PLACE1011978
 PLACE2000118 PLACE3000181 PLACE3000213
 PLACE4000354 SKNMC1000014 SKNMC1000082
 25 THYRO1000061 THYRO1000196 THYRO1000400
 THYRO1000580 THYRO1000584 THYRO1000678
 THYRO1000776 THYRO1000795 THYRO1000846
 THYRO1000866 THYRO1000956 THYRO1000964
 THYRO1001063 THYRO1001071 THYRO1001102
 30 THYRO1001113 THYRO1001128 THYRO1001205
 THYRO1001242 THYRO1001266 THYRO1001456
 THYRO1001457 THYRO1001471 THYRO1001478
 THYRO1001529 THYRO1001593 THYRO1001608
 THYRO1001641 THYRO1001700 THYRO1001702
 35 THYRO1001770 THYRO1001803 Y79AA1000030
 Y79AA1000127 Y79AA1000207 Y79AA1000270
 Y79AA1000426 Y79AA1000750 Y79AA1000777
 Y79AA1000876 Y79AA1000888 Y79AA1000967
 Y79AA1001062 Y79AA1001090 Y79AA1001212
 40 Y79AA1001272 Y79AA1001426 Y79AA1001523
 Y79AA1001727 Y79AA1001787 Y79AA1001799
 Y79AA1001803 Y79AA1001863 Y79AA1002058
 Y79AA1002121 Y79AA1002129 Y79AA1002213
 Y79AA1002334 Y79AA1002376 Y79AA1002378
 45 Y79AA1002381 NT2RP2006580

[0059] Among the clones, the following 83 clones are identical to the clones selected by the score in the ATGpr and the prediction by the PSORT for the existence of a signal sequence.

50 HEMBA1000907 NT2RM2000410 PLACE1000740
 HEMBA1002164 NT2RM2001941 PLACE1001407
 HEMBA1002421 NT2RP1000050 PLACE1002080
 HEMBA1003101 NT2RP2001495 PLACE1002095
 HEMBA1004797 NT2RP2001948 PLACE1002374
 55 HEMBA1006335 NT2RP2002063 PLACE1003428
 HEMBA1006572 NT2RP2002304 PLACE1003460
 HEMBA1006707 NT2RP2003469 PLACE1004482
 HEMBA1006902 NT2RP2003593 PLACE1005383

EP 1 130 094 A2

HEMBA1007013 NT2RP2003655 PLACE1005544
 HEMBB1000447 NT2RP2003664 PLACE1005569
 HEMBB1000567 NT2RP2004447 PLACE1005927
 HEMBB1001200 NT2RP2006042 PLACE1006277
 5 HEMBB1002427 NT2RP2006269 PLACE1006443
 MAMMA1000591 NT2RP3000481 PLACE1007096
 MAMMA1000681 NT2RP3000645 PLACE1007296
 MAMMA1001043 NT2RP3001012 PLACE1008469
 MAMMA1001893 NT2RP3001195 PLACE1008984
 10 MAMMA1001957 NT2RP3001560 PLACE1008985
 MAMMA1002070 NT2RP3002160 PLACE1009527
 MAMMA1002165 NT2RP3002836 PLACE1010251
 MAMMA1002633 NT2RP3002958 PLACE1011236
 NT2RP3003076 SKNMC1000014
 15 NT2RP3003354 THYPO1000196
 NT2RP3004133 THYRO1000795
 NT2RP3004309 THYRO1001478
 NT2RP4001879 THYRO1001702
 NT2RP4002451 Y79AA1000270
 20 OVARC1000439 Y79AA1000426
 OVARC1001222 Y79AA1001803
 Y79AA1001863

25 **[0060]** The 446 clones in Table 16, 17, 18, 19, and 20, and NT2RP2006580 are predicted to encode a secretory or membrane protein. Among them, 77 clones were identical to the clones selected by the score in the ATGpr and the prediction by the PSORT for the existence of a signal sequence (overlapping with any of the 254 clones listed in Table 7-15). Besides, many clones were turned out to be identical to the clones selected as a protein associated with a glycoprotein. Also, there were clones identical to those selected as a protein associated with a disease.

30 **[0061]** The 243 clones in Table 21 are predicted to encode a glycoprotein. Among them, 53 clones were identical to those selected by the score in the ATGpr and the prediction by the PSORT for the existence of a signal sequence. And, many clones were turned out to be identical to the clones selected as a secretory or membrane protein. Moreover, there were clones identical to those selected as a protein associated with a disease.

[0062] The 51 clones in Table 22 are predicted to encode a protein associated to signal transduction.

[0063] The 130 clones in Table 23 are predicted to encode a protein associated to transcription.

35 **[0064]** The 17 clones in Table 24 are predicted to encode a protein associated with diseases.

[0065] In these clones, 532 clones have the maximal ATGpr1 score of 0.5 or higher (Table 25). 60 clones have the maximal ATGpr1 score of 0.3 or higher and less than 0.5 (Table 26 and NT2RP2006580). And 67 clones were with the maximal ATGpr1 score of 0 or higher and less than 0.3 (Table 27).

40 **[0066]** 532 clones shown in Table 25, each having the maximal score in the ATGpr1 0.5 or higher, are very likely to be full-length and encode a secretory or membrane protein, or proteins associated to signal transduction, glycoprotein, transcription, or diseases. 59 clones in Table 26 and NT2RP2006580, which have the maximal score in the ATGpr1 0.3 or higher and less than 0.5, are likely to be full-length and encode a secretory or membrane protein, or proteins associated to signal transduction, glycoprotein, transcription, or diseases. 67 clones in Table 27, having the maximal score in the ATGpr1 0 or higher and less than 0.3, are still likely to be full-length and encode a secretory or membrane protein, or proteins associated to signal transduction, glycoprotein, transcription, or diseases.

45 **[0067]** This is the method for selecting the cDNA clones predicted to encode secretory and/or transmembrane proteins, glycoprotein-associated proteins, signal transduction-associated proteins, transcription-associated proteins, or disease-associated proteins on the basis of the partial sequences (5' sequences).

50 **[0068]** The polynucleotide of the present invention encodes an amino acid sequence of a functional protein such as a secretory or membrane protein, or a protein associated to signal transduction, glycoprotein, transcription, or diseases. Since the protein has the complete amino acid sequence, it is possible to analyze its biological activity by expressing the protein as a recombinant protein using an appropriate expression system, or by raising and using an antibody which specifically recognizes it.

55 **[0069]** It is possible to analyze the biological activity of a secretory protein or a membrane protein, or proteins associated to signal transduction, glycoprotein, or transcription, based on the methods in "Gene Transcription" (Hames B. D., and Higgins S.J. edit, (1993)), "Glycobiology" (Fukuda M., and Kobata A. edit, (1993)), "Growth Factors" (McKay I., and Leigh I. edit, (1993)), "Extracellular Matrix" (Haralson M.A., and Hassell J.R. edit, (1995)), "Transcription Factors" (Latchman D.S. edit, (1993)), "Signal Transduction" (Milligans G. edit, (1992)), featured in "The Practical Approach

EP 1 130 094 A2

Series" (IRL PRESS), or "Signal Transduction Protocols" (Kendall D.A., and Hill S.J. edit, (1995), "Glycoprotein Analysis in Biomedicine" (Hounsell E.F. edit, (1993)), featured in "Method in Molecular Biology" (Humana Press).

[0070] As to a protein associated with a disease, it is possible to perform a functional analysis as described above, but also possible to analyze correlation between the expression or the activity of the protein and a certain disease by using a specific antibody that recognizes the protein. Alternatively, it is possible to utilize the database Online Mendelian Inheritance in Man (OMIM) (<http://www.ncbi.nlm.nih.gov/Omim/>), which is a database of human genes and diseases, to analyze the protein. New information is constantly being deposited in the OMIM database. Therefore, it is possible for one skilled in the art to find a new relationship between a particular disease and a gene of the present invention in the updated database.

[0071] Proteins associated with diseases are useful in drug development as they can be utilized as a diagnostic marker, a drug that regulates the level of their expression and activities, or a target of gene therapy. Also, as for a secretory protein, membrane protein, or proteins associated with signal transduction, glycoprotein, or transcription, search of the OMIM with the keywords mentioned below revealed that the proteins are associated with many diseases. Also, relationship between a proteins associated to signal transduction or transcription and diseases is reported in "Transcription Factor Research-1999" (Fujii, Tamura, Morohashi, Kageyama, and Satake edit, (1999) Jikken-Igaku Zoukan, Vol.17, No.3), and "Gene Medicine" ((1999) Vol.3, No.2). Thus, not only a protein associated with diseases, but also a secretory protein, membrane protein, or protein associated with signal transduction, glycoprotein, or transcription is involved in diseases, suggesting these proteins also are very important as a target in medical industry.

[0072] Keywords used in the search of the OMIM

- (1) secretion protein
- (2) membrane protein
- (3) channel
- (4) extracellular matrix
- (5) receptor
- (6) glycoprotein
- (7) protein kinase
- (8) calmodulin kinase
- (9) transcription factor

[0073] Shown in the search result are only the accession numbers in the OMIM. Using the number, data showing the relationship between a disease and a gene or protein can be seen. The OMIM data has been renewed everyday.

1) Secretion protein

268 entries found, searching for "secretion protein"

104760, 176860, 160900, 107400, 118910, 139320, 603850, 147572, 176880, 600946, 603215, 157147, 600174, 151675, 170280, 179512, 179513, 138120, 179509, 246700, 179510, 600626, 179511, 600998, 109270, 601489, 154545, 179490, 185860, 603216, 122559, 601746, 147290, 602672, 146770, 603062, 179508, 131230, 601591, 602421, 139250, 167805, 167770, 600041, 600564, 118825, 601146, 300090, 600753, 601652, 600759, 600768, 602434, 182590, 603166, 308230, 602534, 603489, 107470, 150390, 104610, 173120, 158106, 143890, 306900, 308700, 134797, 137350, 227500, 176300, 107730, 600760, 138079, 120180, 120160, 120150, 124092, 138160, 101000, 227600, 600509, 601199, 142410, 104311, 193400, 201910, 107300, 122560, 272800, 217000, 590050, 147670, 133170, 176730, 300300, 134370, 274600, 120140, 162151, 158070, 152790, 120120, 106100, 300200, 192340, 190160, 138040, 147470, 147620, 173350, 147380, 152200, 152760, 157145, 153450, 264080, 113811, 600937, 600840, 188545, 202110, 600514, 186590, 603372, 136435, 137241, 252800, 214500, 207750, 138850, 139191, 142640, 138130, 189907, 603692, 600633, 603355, 107270, 600377, 147892, 232200, 600281, 232800, 602358, 137035, 601771, 601769, 253200, 601933, 118444, 600270, 120700, 600945, 603732, 147660, 600761, 172400, 600823, 600877, 130080, 171060, 107740, 307800, 602843, 130660, 152780, 124020, 601124, 601340, 601604, 601610, 171050, 312060, 232700, 300159, 142703, 600734, 125255, 168450, 123812, 188540, 147940, 188450, 600839, 182452, 188400, 182280, 176760, 263200, 600264, 188826, 252650, 601185, 162641, 137216, 601398, 601538, 118888, 118445, 601745, 190180, 601922, 182098, 602008, 147440, 602384, 600031, 109160, 602663, 151670, 602682, 602730, 602779, 146880, 603061, 142704, 603140, 106150, 600732, 153620, 603318, 139392, 600042, 102200, 603493, 182100, 264300, 603795, 184600

2) Membrane protein

1017 entries found, searching for "membrane protein"

130500, 305360, 153330, 173610, 170995, 109270, 170993, 309060, 120920, 602333, 133740, 133710, 602690,

EP 1 130 094 A2

133730, 159430, 600897, 133090, 601178, 602413, 602003, 109280, 603237, 602173, 107776, 602334, 125305, 602335, 182879, 154045, 309845, 600594, 603718, 603241, 603214, 603657, 603177, 600182, 601476, 602879, 136950, 600723, 601114, 185880, 185881, 300096, 602257, 160900, 177070, 603062, 603344, 602977, 310200, 600959, 300100, 186945, 600039, 600267, 128240, 182900, 601097, 136430, 600946, 602534, 601047, 143450, 5 603141, 603700, 600579, 256540, 159440, 602414, 600403, 602048, 188860, 137290, 158343, 184756, 602910, 603179, 600279, 108733, 107770, 173335, 602625, 154050, 219800, 603850, 601028, 600447, 104225, 186946, 601767, 603143, 121015, 603215, 227400, 603735, 600179, 602421, 180721, 176801, 176860, 600753, 603142, 176790, 600266, 601239, 115501, 143890, 121014, 121011, 125950, 603534, 304040, 601134, 600754, 601510, 601595, 190315, 300172, 602216, 602261, 602262, 602461, 131560, 179514, 10 179512, 176981, 142461, 139310, 312080, 176640, 128239, 185470, 310300, 601403, 601757, 273800, 151460, 176943, 104311, 168468, 120130, 602887, 600164, 601531, 601832, 104775, 600040, 603583, 176894, 602631, 166945, 182180, 120620, 141180, 601014, 139150, 182860, 177061, 600174, 180069, 191275, 104760, 601693, 300017, 603518, 601009, 134651, 601107, 103868, 600168, 136425, 603531, 603291, 600917, 603216, 102720, 300118, 179590, 135630, 602285, 107450, 602296, 303630, 176878, 120090, 600322, 138160, 601212, 603293, 15 131230, 112205, 600763, 600718, 300187, 170715, 601966, 300051, 602474, 120070, 600691, 600855, 182309, 602101, 602857, 194355, 162230, 600874, 113730, 155550, 602701, 306400, 601789, 231200, 107271, 175100, 182870, 305100, 301000, 601313, 157147, 147670, 139200, 603593, 157655, 600934, 155970, 602049, 155960, 155760, 118990, 135620, 308230, 602694, 162060, 300023, 160993, 153619, 153432, 120131, 603823, 603167, 601023, 600816, 165040, 601681, 166490, 300112, 120190, 300145, 163970, 20 600772, 602926, 602933, 602202, 400015, 151510, 600759, 602672, 602654, 603821, 116952, 151430, 602632, 155975, 602217, 150370, 600752, 601179, 600932, 603048, 603234, 601805, 603822, 603869, 601717, 601181, 313440, 139130, 107777, 109190, 603452, 191163, 191164, 602370, 176877, 103195, 600523, 191328, 601275, 204200, 602426, 603820, 600551, 600695, 600552, 600553, 602306, 601523, 602507, 602299, 600583, 114070, 600632, 603498, 185430, 600587, 235200, 173470, 603199, 601633, 602500, 25 208900, 180297, 156225, 516020, 190195, 141900, 102680, 193300, 101000, 193400, 300011, 107400, 257220, 107741, 180380, 203200, 111700, 600024, 304800, 600065, 110750, 179605, 113705, 601638, 222900, 120120, 602509, 602469, 600930, 601383, 176261, 602574, 602997, 311770, 131550, 603616, 308700, 603372, 256100, 224100, 276903, 305900, 516000, 131195, 314555, 601567, 603866, 306900, 103390, 186720, 173850, 601050, 602505, 186590, 246530, 602689, 194380, 300041, 162643, 152790, 120150, 600682, 600106, 272750, 188040, 30 602382, 601497, 113811, 182138, 212138, 601309, 109690, 114760, 176805, 601253, 123900, 602581, 189980, 191190, 110700, 600163, 137167, 600580, 601610, 190000, 123825, 603491, 600135, 186591, 173910, 138140, 107266, 120950, 601081, 603690, 244400, 312700, 171060, 601199, 601758, 170500, 277900, 601997, 314850, 601880, 603009, 120220, 603126, 164920, 602934, 164730, 163890, 603434, 107269, 602909, 600877, 256550, 164761, 602872, 120110, 126150, 158070, 266200, 223360, 250800, 269920, 35 252650, 603355, 154582, 138190, 300035, 602640, 227650, 158120, 153700, 182380, 155740, 204500, 603401, 601975, 300135, 136350, 602924, 300167, 185050, 176100, 300189, 151525, 300200, 165180, 230800, 602158, 602676, 603411, 193245, 120325, 601848, 192500, 603102, 147795, 245900, 137060, 147557, 120650, 602377, 307800, 120930, 308100, 142800, 191092, 232300, 173510, 602225, 180470, 190930, 186357, 134638, 600544, 601373, 600509, 600359, 603784, 600395, 600653, 603754, 601597, 601066, 40 600185, 601295, 600978, 205400, 603274, 600418, 600839, 516050, 601691, 601007, 600650, 600308, 603261, 601193, 600004, 600017, 516040, 253800, 276901, 600019, 257200, 108780, 300037, 300104, 300126, 255125, 203300, 300191, 426000, 302060, 304700, 201475, 252010, 193210, 311030, 306250, 248600, 191740, 108360, 131244, 600423, 232200, 191305, 231680, 103320, 190180, 600493, 111200, 226200, 312600, 600170, 111680, 186910, 203100, 600536, 600238, 186830, 186760, 186745, 186711, 106180, 111203, 103180, 182530, 182160, 45 600644, 307030, 192321, 600667, 125647, 179080, 114207, 114860, 176000, 116930, 600748, 173515, 173325, 600377, 171760, 171050, 118425, 170260, 191315, 600798, 600821, 600823, 600444, 600840, 159465, 600857, 158380, 600867, 154360, 152427, 150330, 110900, 147840, 147360, 147280, 146880, 312610, 120940, 142871, 142790, 600937, 142600, 134390, 111250, 600979, 600997, 142460, 186845, 134635, 601017, 139191, 139090, 138850, 601040, 138720, 122561, 131100, 123610, 217070, 100500, 603377, 50 602354, 603302, 603207, 603086, 602188, 602095, 603867, 603842, 603798, 602602, 601194, 602607, 603713, 603681, 601252, 603648, 603646, 603644, 601282, 601284, 603667, 603712, 603594, 601872, 603425, 601843, 603263, 603208, 601411, 603201, 603189, 601463, 603164, 603152, 603087, 602874, 601492, 602893, 602057, 602859, 602746, 603879, 603510, 602458, 603380, 601581, 603765, 603283, 601599, 601733, 601852, 602316, 601615, 601617, 602184, 602894, 603005, 603030, 603861, 602835, 602136, 600153, 600074, 600046, 600023, 55 601625, 516006, 600018, 600016, 516002, 601590, 313475, 313470, 600244, 600528, 601611, 600282, 600327, 601568, 600368, 601730, 601535, 601745, 601929, 300169, 300150, 300132, 601533, 600385, 600464, 600424, 600429, 601756, 601488, 516005, 251100, 516004, 600918, 516003, 602192, 516001, 240500, 600465, 602241, 602243, 230200, 601485, 601478, 601416, 602297, 601459, 601839, 602314,

EP 1 130 094 A2

193065, 193001, 191306, 600504, 601020, 191191, 602372, 190181, 600534, 188380, 186854, 186360, 600530, 185250, 182331, 600535, 182305, 601296, 600582, 600732, 600734, 600742, 600782, 176802, 176266, 600769, 601883, 600864, 601901, 176260, 173490, 600910, 601905, 171890, 600916, 601987, 602679, 162651, 161555, 160994, 602714, 602715, 602724, 602736, 300007, 602783, 275630, 602836, 270200, 602871, 159460, 602876, 154540, 153900, 602890, 601153, 602190, 602905, 153634, 153337, 602914, 152310, 151690, 151625, 602935, 602974, 150325, 602992, 150320, 250790, 603006, 603007, 603008, 150292, 233690, 603046, 150210, 603061, 147940, 603063, 221770, 223100, 603097, 147880, 603118, 147730, 146928, 146630, 142622, 603149, 603150, 603151, 600923, 138981, 138590, 138330, 216950, 603192, 138297, 603202, 601002, 602343, 138230, 136131, 603217, 603220, 134660, 131390, 131235, 603242, 603243, 130130, 602345, 126455, 601123, 126064, 125240, 602359, 603312, 602380, 603318, 123890, 123836, 603356, 603361, 603366, 123830, 179610, 188060, 123620, 120980, 186355, 118510, 114835, 114217, 113810, 603499, 182310, 111740, 109610, 603548, 603564, 108740, 603598, 603613, 107273, 603626, 602518, 179410, 603647, 602515, 603652, 106195, 602573, 178990, 105210, 104615, 167055, 603717, 104614, 603728, 104210, 603749, 603750, 103850, 602608, 603787, 603788, 603796, 173445, 103220, 102910, 102681, 102670, 102642, 603833, 173391, 102576, 102575, 171833, 102573, 101800, 603875, 601108

3) Channel

272 entries found, searching for "channel"
176266, 600724, 170500, 182390, 123825, 114208, 114205, 601784, 114206, 600937, 114204, 603415, 600053, 114209, 114207, 600760, 118425, 601011, 192500, 176261, 600761, 176260, 600359, 600228, 600877, 602235, 300008, 182389, 182391, 601328, 601534, 600504, 602323, 601958, 602780, 602781, 601327, 601012, 600734, 603208, 182392, 603220, 603219, 603888, 600054, 602232, 601745, 603537, 602604, 603796, 302910, 602866, 601013, 602905, 602906, 600163, 152427, 180901, 600702, 600308, 602754, 107776, 602024, 314555, 601949, 600235, 602023, 176263, 600681, 176265, 193245, 603305, 176258, 602983, 601219, 601141, 176267, 602343, 602726, 138253, 176262, 600003, 600397, 602872, 138249, 600843, 600935, 600580, 600845, 602158, 602106, 176264, 300110, 176257, 602717, 603493, 176268, 600932, 602727, 138254, 603652, 300138, 602420, 600570, 600150, 603583, 602345, 603749, 601142, 176256, 600846, 138252, 602982, 603787, 602836, 603788, 602566, 603651, 602421, 100690, 107777, 100725, 100710, 600509, 603061, 154275, 304040, 154276, 180902, 121014, 602368, 139311, 601383, 108745, 601313, 601042, 600131, 186360, 600109, 600229, 600170, 603319, 601485, 118503, 180903, 602076, 124030, 601059, 601212, 601218, 147450, 600855, 600919, 601154, 601157, 171060, 600968, 182139, 131230, 121015, 600421, 113730, 249210, 310500, 600637, 125950, 118800, 156490, 602974, 104610, 121011, 602522, 118504, 300041, 160900, 601382, 602103, 600465, 602014, 600442, 601109, 602481, 277900, 254210, 138247, 164920, 170280, 171050, 128100, 173910, 600884, 123885, 602887, 600232, 180297, 137192, 600304, 138251, 603053, 300103, 603152, 603199, 118511, 118508, 138079, 600983, 182307, 603324, 305990, 603418, 114080, 232200, 600046, 600040, 602403, 603750, 603785, 104210, 600019, 600300, 182860, 603852, 603853, 603855, 516060

4) Extracellular matrix

167 entries found, searching for "extracellular matrix"
603479, 602201, 601418, 601548, 154870, 115437, 602285, 602262, 602261, 134797, 600754, 120361, 116935, 602263, 603320, 601807, 603321, 185250, 185261, 253700, 128239, 120324, 193300, 276901, 308700, 600514, 600261, 602109, 120140, 120150, 147557, 193400, 600536, 188826, 120180, 118661, 120320, 152200, 135821, 112260, 230740, 602090, 155760, 192975, 190182, 602108, 601463, 186745, 600900, 600985, 600758, 602369, 179590, 601211, 600065, 602178, 600262, 182888, 182889, 151510, 182120, 150325, 190181, 150370, 186355, 193065, 165070, 154705, 147559, 146650, 146640, 153619, 175100, 187380, 231050, 188060, 135820, 156790, 130660, 301870, 128240, 600076, 600119, 193210, 600215, 600245, 121010, 150240, 600309, 600491, 222600, 120328, 600564, 600596, 600616, 600700, 600742, 120325, 138297, 600930, 156225, 601028, 601050, 601105, 253800, 601284, 601313, 120280, 310200, 601492, 120250, 601587, 601636, 601652, 601692, 601728, 120220, 601915, 602048, 155120, 310300, 120210, 120165, 120120, 118940, 116930, 602264, 116806, 602366, 120470, 602415, 602428, 602453, 602505, 602574, 603005, 603196, 603221, 603319, 107269, 216550, 103320, 603489, 603551, 603767, 603799, 603842

5) Receptor (including membrane proteins, and also including transcription factors, since nuclear proteins were not excluded in the search)

1606 entries found, searching for "receptor"
600408, 176943, 107770, 601531, 143890, 313700, 162643, 202200, 147670, 191306, 182131, 192321, 138249, 190120, 600264, 162321, 603613, 603614, 602164, 138251, 182134, 138040, 152790, 602393, 133430, 601769, 304800, 147280, 168468, 147730, 155555, 191190, 109760, 122561, 136537, 602034, 147138, 187930, 312861,

EP 1 130 094 A2

107450, 138981, 182135, 603612, 191191, 600933, 601267, 600168, 120650, 182133, 601919, 108962, 162323, 102576, 600022, 603366, 600052, 136430, 601835, 600825, 600979, 102981, 192974, 602779, 601828, 603499, 601972, 191315, 602601, 603611, 602943, 601300, 191316, 108960, 601299, 147810, 602909, 601834, 108961, 602392, 602910, 603372, 106165, 601199, 180901, 131244, 107470, 136435, 146740, 164920, 139191, 138971, 109690, 100690, 138850, 118504, 147781, 146928, 601373, 126449, 104225, 100725, 164342, 104220, 118503, 136538, 137241, 601663, 601805, 306250, 600377, 600337, 600509, 600983, 126110, 118507, 602548, 162010, 601268, 147671, 601284, 113503, 109190, 603826, 154540, 147569, 600956, 138245, 139139, 155541, 137192, 118445, 104250, 300119, 118511, 601159, 100710, 305660, 104210, 114610, 138253, 146929, 602778, 135630, 147140, 146790, 147045, 601998, 158378, 131243, 600041, 118508, 137142, 113995, 138247, 182137, 147811, 182132, 138248, 600414, 600563, 137160, 138032, 601496, 138252, 138254, 601600, 173800, 146760, 179590, 180903, 109635, 300038, 138244, 182139, 600849, 600644, 601296, 602282, 179611, 143090, 100710, 601076, 600380, 170998, 600845, 118505, 118510, 190160, 104260, 137164, 137140, 600456, 147880, 137190, 601156, 601956, 100730, 601194, 602717, 173880, 600042, 180902, 300093, 602423, 602469, 603028, 601437, 146933, 179610, 173510, 600843, 305915, 137143, 603361, 603031, 137141, 136539, 162332, 118509, 155540, 600233, 601109, 601743, 600409, 147139, 131235, 603506, 603455, 151445, 602408, 601681, 603540, 602454, 602490, 600214, 600798, 600232, 602729, 600730, 601122, 120577, 160994, 603248, 602621, 601305, 176946, 118502, 176888, 602167, 138246, 158050, 601470, 603453, 603065, 137163, 137166, 600846, 118495, 138243, 192975, 138060, 602757, 191311, 603652, 120340, 600315, 136425, 118493, 600731, 601502, 601503, 601751, 603651, 602345, 118496, 601598, 602343, 118494, 186880, 134934, 601007, 600946, 186930, 186970, 133171, 308380, 120620, 126452, 190010, 126450, 186810, 180220, 173410, 188070, 131550, 109691, 147679, 180240, 173490, 300034, 264080, 114131, 146661, 147370, 147545, 136352, 190182, 186780, 126453, 110700, 147851, 308385, 126451, 128239, 600018, 601487, 601040, 305990, 601604, 190181, 188545, 162641, 600445, 231200, 136350, 601586, 164770, 600073, 602376, 138491, 118946, 176761, 176806, 601770, 180245, 164870, 602691, 600939, 109610, 192977, 159530, 602997, 164860, 300007, 146710, 601757, 150370, 118444, 600239, 138033, 601167, 108360, 162642, 173850, 603881, 180246, 193210, 173470, 601469, 134935, 182452, 138492, 135620, 601523, 601790, 600221, 109630, 602451, 147265, 602626, 602001, 602069, 601642, 600848, 102776, 180190, 182098, 176886, 602250, 182454, 600222, 600821, 180247, 602730, 134637, 167055, 165196, 165195, 173391, 601426, 600527, 600167, 153618, 601330, 601970, 600785, 600004, 601043, 109135, 109636, 601613, 600804, 601529, 102581, 600936, 126150, 182451, 300107, 190151, 602746, 165070, 170650, 600421, 100997, 602648, 603243, 300035, 305670, 182455, 162651, 109535, 601320, 120920, 603411, 603691, 601535, 102775, 601166, 180381, 176804, 602188, 600051, 601993, 601773, 603755, 600241, 601050, 603659, 600253, 300130, 137026, 602368, 154570, 176945, 602005, 120980, 602583, 102777, 600524, 602848, 173393, 601711, 104221, 176884, 600441, 108361, 600665, 600446, 600799, 600742, 104219, 603819, 600867, 602351, 601978, 600578, 602566, 600551, 601898, 151443, 602642, 300182, 191164, 602356, 603071, 602190, 601522, 182889, 142703, 601576, 100735, 176882, 602969, 182453, 601895, 602886, 602927, 602697, 601070, 120830, 601108, 176891, 176802, 600024, 600552, 601908, 176981, 603510, 151460, 602237, 602545, 162341, 603748, 161565, 600894, 186711, 137161, 601839, 601592, 603507, 602885, 430000, 104222, 601116, 602643, 601896, 601404, 601381, 600180, 600870, 123889, 600206, 600242, 602992, 601380, 602950, 602174, 603159, 602042, 602297, 602836, 602355, 602304, 603822, 147267, 603030, 603821, 602043, 603820, 603317, 601965, 601524, 601974, 603823, 602728, 601741, 137570, 603500, 602592, 153622, 601204, 186990, 601115, 600283, 182180, 600600, 600066, 600282, 176944, 600926, 600895, 300188, 602657, 602756, 600925, 182888, 603808, 603029, 602702, 602464, 603751, 425000, 601910, 151020, 600553, 603195, 602646, 601909, 603692, 136950, 600543, 164320, 300086, 602933, 602993, 137162, 603281, 153243, 161555, 600011, 600144, 600869, 600493, 602346, 600342, 602336, 113955, 602893, 601577, 603749, 602892, 603810, 602853, 600579, 601044, 603809, 186690, 600752, 603304, 600240, 603232, 600868, 602894, 602337, 603143, 109545, 602891, 602890, 602716, 603142, 602510, 603144, 603141, 300090, 602550, 603145, 107741, 107730, 164761, 208900, 107269, 114208, 235200, 158120, 147440, 300200, 161561, 186830, 151520, 131240, 139250, 601667, 153240, 218030, 147680, 186740, 600044, 107940, 153390, 102582, 600978, 601218, 142409, 601156, 601884, 601789, 125950, 171833, 602221, 139320, 147760, 190000, 184757, 601512, 165180, 166490, 168470, 601366, 176000, 191160, 102578, 103320, 112260, 192240, 109091, 603233, 134370, 254210, 173610, 125480, 600957, 118945, 120470, 136351, 600725, 176797, 602887, 173511, 123910, 601459, 252500, 104615, 601528, 600345, 601890, 151390, 602170, 186790, 186960, 600542, 186973, 600423, 600058, 176730, 186945, 104760, 601309, 601500, 300091, 113705, 214500, 602731, 600726, 176885, 164875, 139330, 104640, 147840, 600065, 131242, 164160, 600565, 601211, 151530, 603599, 180860, 190040, 147720, 222100, 223900, 600555, 600994, 186770, 120420, 602004, 601099, 603776, 601212, 132890, 600007, 158105, 603886, 603884, 176830, 162200, 602458, 601603, 107910, 193400, 133100, 160900, 114855, 601498, 601511, 116831, 131390, 107773, 137800, 155750, 147795, 264600, 114207, 244400, 187040,

EP 1 130 094 A2

126455, 164040, 147683, 263200, 188230, 184756, 147574, 179615, 176960, 176947, 147561, 131320, 107265, 147470, 264700, 173335, 124010, 147050, 132810, 230350, 276000, 114130, 155730, 134638, 254150, 122560, 186940, 208150, 188040, 188400, 100600, 601564, 603875, 603849, 601497, 603254, 603180, 602952, 602775, 602584, 602570, 601045, 600632, 602559, 601883, 600618, 601848, 601766, 5 601753, 600554, 600937, 600694, 311360, 302910, 600173, 300139, 300065, 274000, 305100, 300116, 300022, 163800, 187410, 187020, 603818, 105850, 107941, 307800, 601897, 212895, 603598, 603584, 603583, 603454, 603412, 601872, 114204, 266300, 114841, 116897, 152690, 603261, 603130, 603109, 603061, 603055, 603043, 601771, 154275, 600910, 187011, 188060, 600039, 600464, 186355, 121360, 122559, 123000, 253800, 602901, 125255, 131100, 600837, 602772, 600835, 131530, 602695, 137580, 602623, 139130, 249210, 142408, 602457, 10 600179, 142445, 154276, 245480, 142800, 600998, 602354, 601698, 602272, 601053, 602265, 231550, 229300, 147780, 601985, 147796, 226200, 600276, 147830, 150200, 601937, 189906, 601922, 600341, 100650, 601662, 189965, 159440, 190020, 601054, 600764, 600410, 153440, 205400, 601055, 203800, 600576, 600749, 190090, 601767, 600475, 190080, 154550, 184745, 601721, 601712, 600594, 600739, 600738, 180069, 601020, 601679, 162662, 601266, 601676, 601203, 172400, 601463, 600685, 601599, 164009, 15 600611, 601570, 600626, 600201, 600514, 191100, 195002, 600567, 600633, 191030, 190700, 600014, 303700, 308210, 308230, 191170, 269700, 191092, 192130, 600319, 238600, 195000, 600125, 600577, 261680, 600566, 266100, 309900, 266600, 306900, 600536, 194070, 201910, 188860, 188595, 257220, 600038, 202110, 261000, 600363, 600850, 306100, 602229, 603883, 102645, 186921, 186910, 603785, 186852, 104311, 107266, 107270, 603746, 603732, 109270, 602060, 602041, 148180, 151385, 151625, 123830, 20 182530, 123842, 124030, 134660, 153370, 134850, 601023, 182280, 134797, 182125, 134390, 601047, 135820, 601804, 153420, 601801, 139312, 139392, 181031, 601622, 180700, 147570, 603157, 162150, 601589, 121013, 151440, 602919, 151510, 123810, 123890, 131210, 179730, 602663, 179020, 602630, 601231, 601252, 137295, 138190, 138550, 277700, 601295, 139311, 153455, 176311, 176310, 601313, 133170, 141800, 142995, 602358, 143100, 145505, 300079, 168461, 145981, 158070, 106700, 107400, 107720, 601485, 103950, 108330, 165360, 25 300300, 156490, 603492, 602018, 118485, 164951, 164730, 164720, 147620, 603406, 164343, 119530, 161560, 142910, 602414, 601933, 120290, 164013, 164012, 600667, 601534, 601513, 601520, 164015, 164340, 173515, 164975, 302020, 160782, 601501, 165230, 302060, 165120, 601492, 601483, 167040, 167415, 600714, 192968, 600696, 601458, 601447, 173325, 168820, 176820, 601413, 190198, 600489, 173445, 601754, 600481, 300155, 600453, 154230, 173430, 154030, 173360, 181590, 176876, 176970, 178600, 30 164005, 182160, 274500, 179616, 275120, 151990, 300147, 601863, 211750, 275355, 151675, 190170, 179838, 600750, 180200, 180231, 151430, 600753, 601582, 150240, 601584, 601153, 224500, 601154, 601969, 180440, 600596, 600281, 600262, 227400, 601997, 147684, 147678, 600761, 147581, 602014, 602024, 147573, 601150, 159555, 147559, 600945, 191163, 602065, 147558, 601641, 147450, 602096, 602113, 600583, 602154, 600556, 159350, 147245, 186855, 192020, 191510, 146631, 146630, 180950, 601693, 146255, 602260, 238300, 600231, 35 602293, 194540, 602299, 145680, 602311, 145260, 600230, 602350, 239100, 600508, 143055, 601723, 143010, 155960, 601377, 602382, 275350, 600220, 602396, 304050, 600424, 602421, 142461, 602430, 602431, 142460, 246530, 273300, 249100, 602461, 602465, 141850, 601788, 602498, 140210, 181460, 600422, 300127, 600164, 600765, 602582, 138800, 300101, 138720, 138420, 602609, 600157, 138079, 138030, 600156, 601017, 137240, 602671, 602680, 602682, 136530, 182090, 600807, 601826, 602724, 153245, 40 182331, 210250, 600151, 602738, 131340, 131241, 300008, 131195, 600119, 602809, 602839, 602840, 128240, 126850, 126065, 252650, 253200, 188380, 600810, 600396, 123841, 182810, 123835, 600053, 217030, 602921, 602931, 602932, 601900, 600050, 185430, 185520, 121800, 254770, 259700, 222300, 120940, 120700, 120436, 120360, 603033, 600295, 603035, 262600, 603046, 600009, 263400, 120150, 263520, 603126, 590085, 185605, 118960, 118940, 118455, 603149, 603177, 300019, 603181, 118425, 603207, 45 426000, 116955, 116940, 116930, 300088, 603264, 603272, 603274, 116899, 603302, 313650, 266265, 603331, 114835, 603356, 114830, 312760, 603378, 602061, 310400, 603414, 603437, 113730, 310200, 113710, 147557, 113000, 112205, 111400, 111100, 110750, 603518, 603526, 602136, 603576, 309845, 603594, 308310, 108728, 108355, 600250, 307200, 603628, 107777, 602165, 603653, 600899, 107273, 107271, 603695, 603705, 603717, 186590, 186720, 186760, 186820, 187700, 600916, 147183, 105590, 104610, 146770, 306400, 602227, 103390, 50 103280, 103070, 102670, 301000, 603851, 600874, 600871, 102200, 158380, 603889

6) Glycoprotein

438 entries found, searching for "glycoprotein"

231200, 273800, 138700, 138600, 173510, 173515, 192974, 109770, 138610, 603578, 195002, 182889, 601275, 55 602977, 300051, 173470, 171050, 138680, 142640, 194460, 176390, 138720, 600119, 185430, 138670, 180297, 171060, 600073, 173335, 128239, 118850, 602616, 110700, 164345, 159465, 159460, 191845, 173511, 600564, 159440, 256550, 153310, 168810, 600738, 151460, 138470, 190197, 603586, 601325, 601193, 162820, 158120, 155970, 190920, 603201, 603594, 601411, 155735, 186730, 601028, 158030, 176391, 115501, 137290, 601103,

EP 1 130 094 A2

600315, 176870, 601525, 601418, 137061, 137207, 137060, 176397, 176393, 176394, 182888, 160995, 176398, 600718, 176396, 176392, 176399, 176401, 176395, 125645, 601774, 104175, 195000, 601785, 193400, 600065, 111200, 253700, 152200, 154550, 310200, 188230, 256540, 120980, 130660, 107269, 158070, 156790, 131550, 135630, 138297, 155760, 231050, 603381, 143890, 176801, 118860, 103600, 193210, 186820, 114890, 188060, 118910, 107271, 153340, 313470, 600074, 603062, 109270, 602724, 603356, 120324, 192340, 266265, 110900, 306400, 601327, 118457, 151250, 136836, 253000, 104640, 190010, 190000, 135620, 135600, 601852, 134797, 134370, 111250, 156225, 600596, 187430, 164870, 151510, 186720, 601962, 600900, 151530, 180990, 180290, 182120, 603759, 179605, 158381, 158373, 158343, 155740, 153420, 151675, 142800, 176300, 153240, 140100, 134820, 134660, 134390, 125240, 120520, 112205, 111740, 106100, 104225, 103000, 139090, 176860, 102720, 176895, 180621, 602421, 602257, 252600, 252650, 180620, 227600, 228960, 208900, 182145, 208400, 602643, 602593, 602461, 602171, 182900, 601269, 600270, 600033, 602894, 247100, 186910, 600239, 230500, 217800, 600867, 253220, 601581, 305600, 162060, 158378, 188450, 186880, 186990, 188040, 154582, 164022, 188540, 153634, 189965, 193190, 165095, 165220, 158375, 152790, 158374, 151627, 182205, 155750, 191316, 151523, 151020, 165330, 166490, 192225, 224050, 224100, 185490, 192240, 173445, 167770, 185050, 153619, 230800, 147740, 147730, 147670, 153618, 236200, 193300, 249000, 176385, 153440, 147557, 152310, 253270, 184900, 182331, 146760, 146661, 272300, 272800, 142930, 169615, 171640, 182520, 182310, 300027, 147780, 305370, 151525, 142858, 142840, 308230, 308385, 309060, 142445, 142290, 216950, 139130, 600026, 600044, 600046, 138971, 138960, 138140, 600133, 600182, 600215, 173350, 173360, 135820, 600403, 600410, 600491, 134934, 147563, 600535, 133700, 600587, 133170, 131390, 600616, 600713, 131195, 600722, 130160, 600742, 600751, 600764, 600769, 600798, 180540, 600978, 129010, 601040, 601050, 125505, 125490, 124010, 601194, 601211, 601253, 147380, 120930, 601313, 120830, 120620, 601365, 601368, 120220, 601456, 601484, 117700, 147045, 601652, 274500, 116930, 116880, 601788, 116805, 114250, 113810, 601880, 601914, 601932, 111700, 602046, 602053, 602108, 602243, 176820, 178630, 178640, 111680, 300017, 109535, 142860, 602692, 109530, 602746, 308840, 107740, 107400, 603104, 603125, 603126, 603130, 107273, 603197, 107266, 603241, 603243, 104614, 312080, 104610, 603492, 103390, 102670, 102480, 100735, 603665, 603756, 100730, 603777

7) Protein kinase (a member of signal transduction)

729 entries found, searching for "protein kinase"

600289, 160900, 601032, 176948, 600899, 176947, 176894, 601939, 602980, 601314, 601955, 601335, 600447, 602130, 603606, 601158, 602006, 602549, 602678, 176872, 603014, 602520, 186945, 601254, 602315, 601263, 177061, 176873, 600497, 603607, 602449, 602740, 602898, 602741, 300083, 602899, 114080, 603412, 601184, 602739, 176910, 600050, 176949, 602741, 602743, 600341, 601951, 601212, 600038, 156490, 600855, 600758, 603455, 156491, 603424, 300075, 180381, 151510, 600168, 600085, 601890, 602896, 602123, 600221, 602904, 602687, 176945, 602897, 600222, 602098, 601685, 176944, 109635, 601795, 601679, 603722, 602521, 601436, 191164, 602004, 176942, 601684, 300101, 601299, 601434, 602615, 601435, 176960, 153390, 603166, 179090, 300300, 176871, 602216, 176912, 176893, 603453, 300185, 177060, 176943, 176981, 603441, 176970, 116899, 603442, 311550, 603495, 169190, 600448, 602399, 176946, 176977, 188830, 602350, 601591, 601782, 137026, 400008, 300094, 600058, 600831, 176892, 601639, 602919, 176911, 176975, 300172, 603072, 600583, 177015, 176982, 176980, 603248, 176941, 602757, 600870, 603440, 148750, 601959, 600869, 602448, 600408, 306000, 266200, 602539, 602425, 600160, 310200, 208900, 600997, 186973, 305360, 179611, 307030, 603262, 601050, 602758, 147795, 311800, 600664, 600441, 603251, 600527, 313650, 123831, 191315, 602958, 123832, 602989, 602614, 123828, 136350, 601284, 602940, 601329, 600173, 171834, 124095, 603464, 603584, 116953, 300035, 600600, 109135, 601232, 601988, 603850, 603384, 603184, 602990, 125855, 601047, 603460, 603258, 602748, 603723, 164870, 603368, 139310, 601818, 602913, 603005, 602221, 173570, 147670, 602610, 600524, 113508, 251170, 142370, 142408, 191306, 147796, 118423, 600286, 601074, 600765, 600922, 179610, 602747, 601953, 601132, 600136, 601288, 601289, 602297, 191316, 601535, 602925, 125450, 602839, 602168, 602461, 105590, 114085, 600714, 154235, 603261, 109270, 191305, 602625, 602366, 602749, 600927, 300134, 603496, 136352, 176790, 600073, 601523, 115441, 191170, 114210, 116900, 188250, 102576, 603764, 603497, 603304, 601934, 602255, 600004, 600526, 601589, 600856, 601269, 601465, 601207, 601935, 600863, 602394, 601983, 602048, 300200, 179050, 164761, 189980, 602626, 137025, 602694, 602731, 125305, 123810, 300139, 601792, 275350, 171833, 193525, 601540, 162200, 601826, 116898, 601108, 601528, 600954, 602609, 602838, 603167, 601522, 600505, 600431, 600267, 188555, 600140, 308240, 600137, 600011, 151410, 601530, 134934, 116952, 300142, 191311, 601603, 603889, 601728, 165160, 165070, 136351, 601836, 601839, 133090, 601595, 602745, 602516, 154950, 603601, 602052, 154045, 603583, 600695, 602933, 602756, 601014, 602474, 602887, 172270, 601366, 180220, 602189, 114105, 600963, 603369, 601296, 603289, 602265, 602337, 600066, 600456, 602387, 600917, 139270, 603271, 188345, 601231, 115440, 115442, 600778, 164920, 175100, 310300, 601767, 603168, 603140, 603113,

EP 1 130 094 A2

116940, 162060, 603048, 603275, 601146, 600434, 602336, 602310, 602308, 147522, 603311, 227645, 114110, 601114, 300022, 164785, 167414, 123280, 600007, 600560, 229300, 151430, 176740, 602426, 109636, 602775, 108355, 601999, 601524, 190151, 202500, 601592, 600723, 102750, 600753, 603618, 603814, 147521, 602514, 603015, 602525, 602524, 172471, 602647, 602590, 600864, 172470, 602538, 5 602527, 602526, 143890, 164730, 232300, 131550, 116951, 601713, 154550, 603434, 138160, 601702, 126335, 272800, 601497, 602038, 107940, 191190, 273800, 193300, 120150, 603068, 603259, 102582, 601099, 253700, 308000, 602338, 602689, 185605, 164040, 600119, 152690, 601496, 602188, 123900, 602860, 164960, 600618, 601045, 300189, 164860, 602505, 230800, 141850, 147545, 110300, 600051, 138600, 233700, 109700, 238600, 109690, 253800, 311770, 156225, 603719, 191030, 309000, 601893, 190030, 189972, 601247, 190120, 600179, 10 258501, 190040, 300121, 180380, 314850, 600438, 306400, 179715, 190450, 600636, 180901, 173335, 601290, 154500, 131195, 115460, 603078, 603775, 603760, 158900, 602354, 102680, 122560, 602382, 103320, 603597, 603510, 232600, 133170, 603353, 603345, 603318, 603309, 230400, 139250, 603296, 603272, 114240, 147557, 232400, 139259, 602122, 230200, 164014, 248600, 601703, 194364, 150000, 602102, 601683, 158070, 602170, 602007, 601306, 248610, 601387, 161770, 132800, 164760, 601902, 601253, 15 173470, 602930, 603691, 165360, 123835, 603544, 147265, 107777, 602377, 600605, 602869, 601121, 602388, 301300, 590060, 601059, 168461, 142600, 314555, 142410, 170290, 602498, 190090, 311030, 142409, 190181, 600966, 308100, 600936, 171900, 190182, 172405, 102600, 603754, 600281, 172430, 138249, 300039, 603154, 602642, 300041, 602659, 192240, 603473, 136537, 603134, 182280, 603084, 125660, 189990, 133430, 603755, 603302, 600737, 300144, 305400, 188545, 600652, 189940, 600633, 603009, 20 181031, 182160, 148500, 602384, 600650, 602190, 171150, 173490, 164765, 147780, 600910, 601272, 147880, 125255, 120090, 173390, 602668, 600483, 602995, 602996, 180860, 142445, 116949, 602681, 603027, 182125, 600327, 600799, 164757, 230450, 602730, 600782, 164880, 232700, 116935, 602303, 603203, 164710, 600244, 600206, 263200, 300034, 601487, 600980, 600956, 601985, 600884, 601956, 191041, 601568, 601586, 155960, 602288, 176761, 113900, 600581, 516050, 190070, 164940, 600544, 107941, 312861, 602373, 123836, 106165, 25 103850, 274270, 601922, 602355, 147420, 311010, 163800, 600536, 602018, 307200, 603619, 603652, 193001, 193245, 603729, 603752, 118888, 603796, 305600, 603852, 603853, 603886

8) Calmodulin kinase (a member of signal transduction)

35 entries found, searching for "calmodulin binding"
30 300172, 114180, 302020, 139312, 602584, 602293, 600310, 603379, 114106, 602350, 114105, 114213, 138249, 153430, 177061, 182520, 102680, 600489, 601478, 306000, 600490, 600922, 601302, 601568, 168890, 164795, 163730, 602698, 602699, 602987, 102681, 603384

9) Transcription factor

717 entries found, searching for "transcription factor"
35 305371, 189963, 164177, 600297, 600298, 163260, 189907, 173110, 600733, 189903, 189972, 600438, 600281, 189962, 157670, 602272, 600609, 189968, 189971, 189889, 107580, 189908, 601748, 601542, 601632, 600610, 189969, 600480, 601760, 176310, 603022, 600519, 601750, 123803, 600520, 600494, 164176, 602460, 603246, 113725, 107773, 189904, 602228, 147141, 132890, 189964, 164175, 142410, 313650, 164011, 164014, 164005, 40 601714, 600492, 164012, 601878, 600673, 600635, 191523, 602191, 600490, 602149, 189967, 140580, 164343, 603256, 600727, 600729, 600728, 140581, 600571, 601511, 102582, 189906, 600660, 124097, 314310, 600860, 600662, 184756, 600777, 600013, 600172, 600489, 601425, 600786, 600502, 602406, 600743, 189965, 603257, 602751, 602542, 603148, 603107, 603789, 602318, 156845, 123811, 600425, 600540, 600695, 601622, 147574, 601601, 602407, 602150, 601538, 600663, 603306, 165170, 187040, 602444, 45 189902, 189973, 600659, 600661, 601010, 600788, 602617, 601602, 602053, 601742, 300039, 602438, 602976, 600744, 602543, 602479, 600481, 600473, 603739, 600426, 603738, 189901, 603677, 600427, 603255, 600607, 600379, 189909, 601679, 600787, 602160, 601043, 601397, 601044, 600366, 300025, 602575, 602669, 601804, 601801, 176312, 176311, 600746, 602480, 602944, 600967, 600912, 306700, 306900, 193400, 601206, 480000, 191160, 601861, 164008, 600475, 600773, 600772, 600774, 142409, 156490, 600589, 601490, 151385, 600599, 50 184757, 602955, 234000, 603433, 603349, 603198, 602294, 600390, 603628, 147620, 600211, 601787, 601863, 147470, 603795, 603734, 152760, 104155, 128990, 601729, 600197, 147370, 173490, 603423, 600822, 188595, 603243, 600573, 601689, 142765, 603181, 600879, 603731, 600288, 602295, 121360, 164874, 300019, 162095, 602355, 603258, 126090, 159540, 300070, 600555, 600664, 601874, 153245, 191191, 601126, 601512, 146733, 131550, 142385, 601796, 603406, 602959, 601734, 601732, 139191, 139139, 55 600633, 138971, 600006, 603170, 601488, 147576, 147680, 601498, 602630, 602643, 603364, 600914, 154040, 602746, 128992, 143089, 160900, 600140, 134934, 133510, 176860, 190180, 601150, 601175, 170993, 601361, 122560, 602778, 308230, 602903, 309550, 601788, 602946, 159970, 124092, 180200, 173410, 602356, 603015, 600779, 603111, 187930, 602614, 600951, 603200, 602369, 164770, 147569, 603300, 603301, 159980, 134638,

603431, 147730, 603366, 603348, 600556, 602136, 164160, 310200, 152390,
 601241, 116897, 137295, 600576, 194070, 601487, 600698, 164810, 601769, 141900, 602225, 275350, 131100,
 179755, 600075, 162200, 165160, 116806, 600899, 123810, 133450, 216400, 278700, 190080, 164730, 191170,
 193300, 600618, 600999, 601090, 106150, 601843, 133530, 110700, 602550, 138040, 133430, 300133, 163731,
 5 602302, 126337, 309548, 180245, 126110, 602291, 109565, 107400, 314670, 601444, 143100, 104760, 106180,
 601953, 600584, 125852, 602419, 600401, 142200, 107680, 167414, 600020, 188400, 208900, 175100, 602700,
 601828, 139320, 602777, 600185, 602681, 603023, 314997, 602848, 600284, 102578, 114290, 165095, 137070,
 602991, 602421, 600005, 602996, 314995, 152200, 151900, 112260, 129010, 600892, 273800, 176760, 602341,
 490000, 136533, 400003, 601007, 602229, 603620, 602218, 602116, 602020,
 10 142000, 601955, 126340, 120150, 193067, 182452, 142461, 194558, 180660, 600756, 160745, 107741, 106210,
 157640, 186770, 146738, 603759, 213700, 147880, 152391, 277900, 151626, 107730, 600711, 600246, 107470,
 600237, 223100, 107748, 600065, 600349, 426000, 194500, 154030, 227650, 600247, 314980, 109560, 305370,
 600800, 301000, 277700, 600838, 312173, 600439, 600440, 191315, 601595, 190450, 190070, 190020, 162360,
 131320, 133540, 600993, 601993, 159530, 601902, 602868, 181590, 601724, 602260, 601093, 187270, 164761,
 15 602102, 603245, 136950, 106100, 601182, 167410, 601897, 602896, 170100, 602506, 104150, 176730, 601600,
 187011, 102600, 180380, 162080, 603450, 142967, 602301, 126375, 603372, 603355, 164720, 603250, 167409,
 167415, 602897, 601565, 185250, 182138, 601851, 600749, 601575, 194548,
 154500, 601365, 194541, 601621, 601623, 601531, 600790, 194355, 123830, 123812, 154540, 601415, 143055,
 601386, 194550, 186930, 131290, 601320, 601620, 601754, 601313, 184430, 182900, 182500, 600725, 147870,
 20 154365, 116953, 601297, 601296, 601265, 600796, 120436, 601644, 601930, 601643, 230200, 601645, 601972,
 600861, 602009, 601172, 601158, 601646, 180630, 600821, 118440, 601656, 601647, 150200, 601125, 601671,
 141850, 116899, 600697, 109270, 202110, 150570, 601108, 191339, 601063, 109691, 180240, 203100, 151430,
 179710, 111000, 176797, 238600, 104311, 240300, 125255, 600423, 158070, 602439, 600324, 112261, 243305,
 602474, 174762, 600613, 602539, 138890, 138720, 114550, 173865, 602582, 602584, 173510, 600250, 602627,
 25 173325, 602635, 246530, 172425, 600193, 602691, 600188, 170998, 152790,
 168468, 256540, 225250, 600848, 143400, 168461, 262600, 168360, 601912, 602951, 600017, 230000, 266600,
 602981, 272800, 109150, 102200, 603025, 603026, 603109, 167050, 603127, 603128, 165240, 230400, 313700,
 164975, 164875, 602017, 115500, 235800, 164873, 602110, 164785, 164772, 312865, 603296, 600542, 164740,
 602125, 309801, 602148, 300007, 306955, 603368, 116940, 602181, 603416, 126650, 163920, 300024, 603437,
 30 602209, 603576, 603607, 305435, 600944, 180410, 303630, 159557, 301870, 132810, 100790, 603849, 603862,
 603881,

[0074] There are several methods for analyzing the expression levels of genes associated with diseases. Differences
 in gene expression levels between diseased and normal tissues are studied by the analytical methods, for example,
 35 Northern hybridization and differential display. Other examples include a method with high-density cDNA filter, a method
 with DNA microarray and methods with PCR amplification (Experimental Medicine, Vol.17, No. 8, 980-1056 (1999);
 Cell Engineering (additional volume) DNA Microarray and Advanced PCR Methods, Muramatsu & Naba (eds.), Shu-
 junsya). The levels of gene expression between diseased tissues and normal tissues can be studied by any of these
 analytical methods. When explicit difference in expression level is observed for a gene, it can be concluded that the
 40 gene is closely associated with a disease or disorder. Instead of diseased tissues, cultured cells can be used for the
 assessment. Similarly, when gene expression is explicitly different between normal cells and cells reproducing disease-
 associated specific features, it can be concluded that the gene is closely associated with a disease or disorder. When
 the expression levels of genes are evidently varied during major cellular events (such as differentiation and apoptosis),
 the genes are involved in the cellular events and accordingly are candidates for disease- and/or disorder-associated
 45 genes. Further, genes exhibiting tissue-specific expression are genes playing important parts in the tissue functions
 and, therefore, can be candidates for genes associated with diseases and/or disorders affecting the tissues.

[0075] For example, non-enzymic protein glycation reaction is believed to be a cause for a variety of chronic diabetic
 complications. Accordingly, in endothelial cells, genes, of which expression levels are elevated or decreased in a gly-
 cated protein-dependent manner, are associated with diabetic complications caused by glycated proteins (Diabetes
 50 1996, 45 (Suppl. 3), S67-S72; Diabetes 1997, 46 (Suppl. 2), S19-S25). The onset of rheumatoid arthritis is thought to
 be involved in the proliferation of synovial cells covering inner surfaces of joint cavity and in inflammatory reaction
 resulted from the action of cytokines produced by leukocytes infiltrating into the joint synovial tissues (Rheumatism
 Information Center, <http://www.rheuma-net.or.jp/>). Recent studies have also revealed that tissue necrosis factor (TNF)-
 α participates in the onset (Current opinion in immunology 1959, 11, 657-662). When the expression of a gene exhibits
 55 responsiveness to the action of TNF on synovial cells, the gene is considered to be involved in rheumatoid arthritis.
 Many genes acting at the downstream of TNF- α and IL-1, among inflammation-associated cytokines have been previ-
 ously identified. The respective stimulations are transduced through independent pathways of signaling cascade. There
 exists another signaling cascade for both stimulations, wherein NF- κ B is a common transducing molecule shared by

the two stimulations (J. Leukoc. Biol., 1994, 56(5): 542-547). It has also been revealed that many inflammation-associated genes, including IL-2, IL-6 and G-CSF, are varied in the expression levels in response to the signal through the common pathway (Trends Genet. 1999, 15(6): 229-235). It is assumed that genes of which expression levels are varied in response to the stimulation of TNF- α or IL-1, also participate in inflammation. Genes associated with neural differentiation can be candidates for causative genes for neurological diseases as well as candidates for genes usable for treating the diseases.

[0076] Clones exhibiting differences in the expression levels thereof can be selected by using gene expression analysis. The selection comprises, for example; analyzing cDNA clones by using high-density cDNA filter; and statistically treating the multiple signal values (signal values of radioisotope in the radiolabeled probes or values obtained by measuring fluorescence intensities emitted from the fluorescent labels) for the respective clones by two-sample t-test, where the signal values are determined by multiple experiments of hybridization. The clones of interest are selectable based on the statistically significant differences in the signal distribution at $p < 0.05$. However, selectable clones with significant difference in the expression levels thereof may be changed depending on the partial modification of statistical treatment. For example, the clones may be selected by conducting statistical treatment with two-sample t-test at $p < 0.01$; or genes exhibiting more explicit differences in the expression levels thereof can be selected by performing statistical treatment with a pre-determined cut-off value for the significant signal difference. An alternative method is that the expression levels are simply compared with each other, and then, the clones of interest are selected based on the ratio of the expression levels thereof.

[0077] Clones exhibiting differences in the expression levels can also be selected by comparing the expression levels by PCR analysis, for example, by using the method of determining the band intensities representing the amounts of PCR products with ethidium bromide staining; the method of determining the radioisotopic signal values or fluorescence intensities of the PCR products when radio-labeled or fluorescence-labeled primers; or the method of determining the values of radioisotope signals or fluorescence intensities of the probes hybridized to the PCR products when radio-labeled or fluorescence-labeled probes, respectively, are used in the hybridization. If the expression level ratios obtained in multiple PCR experiments are constantly at least 2-fold, such a clone can be judged to exhibit the difference in the expression level. When the ratios are several-fold or not less than 10-fold, the clone can be selected as a gene exhibiting the explicit difference in the expression level.

[0078] A survey of genes of which expression levels are varied specifically to the glycosylated protein in the endothelial cells revealed three genes with elevated expression levels, NT2RP2001538, NT2RP4001001 and Y79AA1000967. These clones are genes associated with diabetes.

[0079] A survey of genes of which expression levels are varied in response to TNF. (Tumor Necrosis Factor- α) in the primary cell culture of synovial tissue detected the following clones with elevated expression levels in the presence of TNF- α :

BNGH41000020,	HEMBA1000349,	HEMBA1000634,	HEMBA1000671,	HEMBA1000835,	HEMBA1000962,
HEMBA1002178,	HEMBA1002195,	HEMBA1002239,	HEMBA1002420,	HEMBA1002524,	HEMBA1002992,
HEMBA1003315,	HEMBA1003392,	HEMBA1003487,	HEMBA1003602,	HEMBA1004067,	HEMBA1004797,
HEMBA1005337,	HEMBA1005489,	HEMBA1006916,	HEMBA1006668,	HEMBA1000905,	HEMBA1001547,
HEMBA1001573,	HEMBA1002041,	HEMBA1002663,	MAMMA1000652,	MAMMA1000810,	MAMMA1001634,
MAMMA1002091,	MAMMA1002234,	NT2RM2000306,	NT2RM4000417,	NT2RP1000002,	NT2RP1000181,
NT2RP1000740,	NT2RP2000694,	NT2RP2001921,	NT2RP2002527,	NT2RP2004495,	NT2RP2004606,
NT2RP2005163,	NT2RP2005463,	NT2RP2006134,	NT2RP3000171,	NT2RP3000652,	NT2RP3001195,
NT2RP3001976,	NT2RP3003473,	NT2RP3003874,	NT2RP3004090,	NT2RP3004294,	NT2RP3004557,
NT2RP3004647,	NT2RP4000108,	NT2RP4001001,	NT2RP4001877,	OVARC1000090,	OVARC1000105,
OVARC1000275,	OVARC1000439,	OVARC1001607,	PLACE1000740,	PLACE1000927,	PLACE1001016,
PLACE1001100,	PLACE1001464,	PLACE1001500,	PLACE1001918,	PLACE1002095,	PLACE1002547,
PLACE1003644,	PLACE1004519,	PLACE1005031,	PLACE1005410,	PLACE1005736,	PLACE1006219,
PLACE1006809,	PLACE1008716,	PLACE1010081,	THYRO1001770,	Y79AA1000127,	Y79M1000207,
Y79AA1000270,	Y79AA1000876,	Y79AA1001013,	Y79AA1001264,	Y79AA1001272,	Y79AA1001328,
Y79AA1001430,	Y79AA1001530,	Y79AA1001799			

[0080] Clones with decreased expression levels in the presence of TNF- α are NT2RM4000326, NT2RP1000300, NT2RP2000514, NT2RP2001755, NT2RP2006042, NT2RP3000481, NT2RP3002790. These clones are candidates for rheumatoid arthritis-associated genes.

[0081] A survey of genes of which expression levels are varied in response to TNF. (Tumor Necrosis Factor- α) or IL-1. (Interleukin-1 beta) in a human T cell strain, Jurkat cell, revealed the following clones with elevated expression levels in the presence of TNF- α :

MAMMA1000141,	MAMMA1000788,	MAMMA1001237,	MAMMA1002070,	NT2RM2000582,	NT2RM2002109,
NT2RP1000679,	NT2RP2003664,	NT2RP2004108,	NT2RP2005597,	NT2RP3001592,	NT2RP3002738,
NT2RP3004133,	NT2RP3004294,	NT2RP3004321,	NT2RP3004557,	PLACE1002547,	PLACE1003573,

EP 1 130 094 A2

PLACE1004305, PLACE1008744, PLACE1010011, PLACE1010713, PLACE1011181, Y79AA1000776, Y79AA1002129

[0082] The survey also revealed a clone of which expression level was decreased in the presence of TNF. The clone is PLACE1002070. The same survey further revealed the clones of which expression levels were elevated in the presence of IL-1. The clones are MAMMA1000614, MAMMA1001237, NT2RM2000514 and NT2RP3001159. These clones are genes associated with inflammation.

[0083] A survey of genes of which expression levels are varied in response to the stimulation for inducing cell differentiation (stimulation using retinoic acid (RA) or using RA/inhibitor (inhibitor for cell division)) in cultured cells of neural strain, NT2, revealed the following clones with elevated expression levels in the presence of RA:

10 HEMBA1000121, HEMBA1000275, HEMBA1000300, HEMBA1000634, HEMBA1000671, HEMBA1000875, HEMBA1001184, HEMBA1001390, HEMBA1001886, HEMBA1002163, HEMBA1002227, HEMBA1002420, HEMBA1002421, HEMBA1003072, HEMBA1003120, HEMBA1003294, HEMBA1003497, HEMBA1004007, HEMBA1004110, HEMBA1004391, HEMBA1004444, HEMBA1005230, HEMBA1005246, HEMBA1005267, HEMBA1005489, HEMBA1005913, HEMBA1006299, HEMBA1006357, HEMBA1006517, HEMBA1006544, 15 HEMBA1006658, HEMBA1006749, HEMBA1007063, HEMBA1007241, HEMBB1000447, HEMBB1000542, HEMBB1000567, HEMBB1000642, HEMBB1000668, HEMBB1001026, HEMBB1001847, HEMBB1002051, HEMBB1002120, HEMBB1002228, HEMBB1002693, MAMMA1000106, MAMMA1000141, MAMMA1000473, MAMMA1000528, MAMMA1000810, MAMMA1000881, MAMMA1001634, MAMMA1001957, MAMMA1002205, MAMMA1002224, NT2RM2000423, NT2RM2000497, NT2RM2000582, NT2RM2001126, NT2RM2001902, 20 NT2RM4000198, NT2RM4000284, NT2RM4000593, NT2RM4001321, NT2RP1000002, NT2RP1000050, NT2RP1000181, NT2RP1000261, NT2RP1000465, NT2RP1000468, NT2RP1000579, NT2RP1000679, NT2RP2000092, NT2RP2000479, NT2RP2000610, NT2RP2000663, NT2RP2000694, NT2RP2000903, NT2RP2001388, NT2RP2001538, NT2RP2001878, NT2RP2002015, NT2RP2002304, NT2RP2002721, NT2RP2002824, NT2RP2002942, NT2RP2002974, NT2RP2002976, NT2RP2003179, NT2RP2003302, 25 NT2RP2003383, NT2RP2003469, NT2RP2003664, NT2RP2003940, NT2RP2004069, NT2RP2004108, NT2RP2004524, NT2RP2004556, NT2RP2004670, NT2RP2005069, NT2RP2005247, NT2RP2005425, NT2RP2005463, NT2RP2005514, NT2RP2005535, NT2RP2005541, NT2RP2005774, NT2RP2005878, NT2RP2005883, NT2RP2005887, NT2RP2006099, NT2RP2006134, NT2RP3000011, NT2RP3000125, NT2RP3000171, NT2RP3000232, NT2RP3000460, NT2RP3000481, 30 NT2RP3000652, NT2RP3000677, NT2RP3000818, NT2RP3000820, NT2RP3001044, NT2RP3001061, NT2RP3001170, NT2RP3001240, NT2RP3001322, NT2RP3001388, NT2RP3001542, NT2RP3001592, NT2RP3001976, NT2RP3002790, NT2RP3002900, NT2RP3002983, NT2RP3003000, NT2RP3003354, NT2RP3003532, NT2RP3003729, NT2RP3003874, NT2RP3003939, NT2RP3004025, NT2RP3004083, NT2RP3004090, NT2RP3004130, NT2RP3004202, NT2RP3004294, NT2RP3004640, NT2RP4000108, 35 NT2RP4000634, NT2RP4002451, NT2RP4002715, OVARC1000090, OVARC1000208, OVARC1000275, OVARC1000553, OVARC1000775, OVARC1000853, OVARC1000873, OVARC1000916, OVARC1000995, OVARC1001030, OVARC1001049, OVARC1001132, OVARC1001596, OVARC1002178, OVARC100258, PLACE1000442, PLACE1000927, PLACE1000986, PLACE1001100, PLACE1001123, PLACE1001795, PLACE1002518, PLACE1002547, PLACE1002967, PLACE1003407, PLACE1003428, PLACE1003644, 40 PLACE1003839, PLACE1004078, PLACE1004441, PLACE1004450, PLACE1005669, PLACE1005682, PLACE1005736, PLACE1005768, PLACE1005815, PLACE1006073, PLACE1006208, PLACE1007296, PLACE1007626, PLACE1008282, PLACE1008984, PLACE1008985, PLACE1010445, PLACE1011708, PLACE1011978, PLACE4000455, SKNMC1000004, THYRO1000036, THYRO1000580, THYRO1000776, THYRO1000999, THYRO1001063, THYRO1001128, THYRO1001205, 45 THYRO1001327, THYRO1001523, THYRO1001725, THYRO1001770, Y79AA1000207, Y79AA1000226, Y79AA1000270, Y79AA1001056, Y79AA1001062, Y79AA1001090, Y79AA1001727, Y79AA1002213, Y79AA1002381

[0084] The survey also revealed the clones of which expression levels were decreased in the presence of RA. The clones are BNGH41000020, HEMBA1005070, NT2RP2005027, NT2RP3003473 and Y79AA1002376.

50 **[0085]** The same survey further revealed the following clones with elevated expression levels in the presence of RA/inhibitor:

55 HEMBA1000128, HEMBA1000875, HEMBA1001390, HEMBA1002163, HEMBA1002227, HEMBA1002421, HEMBA1004391, HEMBA1004454, HEMBA1004785, HEMBA1005913, HEMBA1006171, HEMBA1006299, HEMBA1006335, HEMBA1006544, HEMBA1007241, HEMBB1000447, HEMBB1000668, MAMMA1000994, MAMMA1001344, NT2RM2000582, NT2RP1001004, NT2RP2000663, NT2RP2000694, NT2RP2000903, NT2RP2001388, NT2RP2002674, NT2RP2002974, NT2RP2003383, NT2RP2004069, NT2RP2004606, NT2RP2004837, NT2RP2005069, NT2RP2005425, NT2RP2005463, NT2RP2005541, NT2RP2005883, NT2RP2005887, NT2RP3000460, NT2RP3000838, NT2RP3001044, NT2RP3001240, NT2RP3001388,

EP 1 130 094 A2

NT2RP3002721, NT2RP3002738, NT2RP3003469, NT2RP3004083, NT2RP3004130, NT2RP3004202,
 NT2RP3004294, NT2RP3004640, NT2RP4000108, NT2RP4002451, NT2RP4002715, OVARC1000275,
 OVARC1000467, OVARC1000553, OVARC1000853, OVARC1000873, OVARC1000916, OVARC1000995,
 OVARC1001030, OVARC1001222, OVARC1001596, OVARC1002058, OVARC1002178, PLACE1000927,
 5 PLACE1001123, PLACE1001407, PLACE1001464, PLACE1001564, PLACE1001795, PLACE1002547,
 PLACE1003407, PLACE1003644, PLACE1003845, PLACE1004441, PLACE1004482, PLACE1005410,
 PLACE1005601, PLACE1005725, PLACE1005736, PLACE1006093, PLACE1006219, PLACE1006290,
 PLACE1006716, PLACE1007296, PLACE1007626, PLACE1008359, PLACE1010968, PLACE1011364,
 PLACE1011824, THYRO1000678, THYRO1000776, THYRO1000999, THYRO1001113, THYRO1001237,
 10 THYRO1001523, Y79AA1000226, Y79AA1000888, Y79AA1001430

[0086] The same survey further revealed the following clones with elevated expression levels in the presence of RA/ inhibitor:

HEMBA1000349, HEMBA1001297, HEMBA1001878, HEMBA1005070, HEMBA1006482, HEMBB1001959,
 NT2RM2001939, NT2RP1000981, NT2RP2001469, NT2RP3003473, OVARC1001132, PLACE1001655,
 15 Y79AA1000127, Y79AA1002381. These clones are associated with neural differentiation and, therefore, are candidates for genes associated with neurological diseases.

[0087] Based on the functional analyses using a secretory protein, membrane protein, or proteins associated with signal transduction, glycoprotein, transcription, or diseases, it is possible to develop a medicine.

[0088] In case of a membrane protein, it is most likely to be a protein that functions as a receptor or ligand on the cell surface. Therefore, it is possible to reveal a new relationship between a ligand and receptor by screening the membrane protein of the invention based on the binding activity with the known ligand or receptor. Screening can be performed according to the known methods.

[0089] For example, a ligand against the protein of the invention can be screened in the following manner. Namely, a ligand that binds to a specific protein can be screened by a method comprising the steps of: (a) contacting a test sample with the protein of the invention or a partial peptide thereof, or cells expressing these, and (b) selecting a test sample that binds to said protein, said partial peptide, or said cells.

[0090] On the other hand, for example, screening using cells expressing the protein of the present invention that is a receptor protein can also be performed as follows. It is possible to screen receptors that is capable of binding to a specific protein by using procedures (a) attaching the sample cells to the protein of the invention or its partial peptide, and (b) selecting cells that can bind to the said protein or its partial peptide.

[0091] In a following screening as an example, first the protein of the invention is expressed, and the recombinant protein is purified. Next, the purified protein is labeled, binding assay is performed using a various cell lines or primary cultured cells, and cells that are expressing a receptor are selected (Growth and differentiation factors and their receptors, Shin-Seikagaku Jikken Kouza Vol.7 (1991) Honjyo, Arai, Taniguchi, and Muramatsu edit, p203-236, Tokyo-Kagaku-Doujin). A protein of the invention can be labeled with RI such as ¹²⁵I, and enzyme (alkaline phosphatase etc.). Alternatively, a protein of the invention may be used without labeling and then detected by using a labeled antibody against the protein. The cells that are selected by the above screening methods, which express a receptor of the protein of the invention, can be used for the further screening of an agonists or antagonists of the said receptor.

[0092] Once the ligand binding to the protein of the invention, the receptor of the protein of the invention or the cells expressing the receptor are obtained by screening, it is possible to screen a compound that binds to the ligand and receptor. Also it is possible to screen a compound that can inhibit both bindings (agonists or antagonists of the receptor, for example) by utilizing the binding activities.

[0093] When the protein of the invention is a receptor, the screening method comprises the steps of (a) contacting the protein of the invention or cells expressing the protein of the invention with the ligand, in the presence of a test sample, (b) detecting the binding activity between said protein or cells expressing said protein and the ligand, and (c) selecting a compound that reduces said binding activity when compared to the activity in the absence of the test sample. Furthermore, when the protein of the invention is a ligand, the screening method comprises the steps of (a) contacting the protein of the invention with its receptor or cells expressing the receptor in the presence of samples, (b) detecting the binding activity between the protein and its receptor or the cells expressing the receptor, and (c) selecting a compound that can potentially reduce the binding activity compared to the activity in the absence of the sample.

[0094] Samples to screen include cell extracts, expressed products from a gene library, synthesized low molecular compound, synthesized peptide, and natural compounds, for example, but are not construed to be listed here. A compound that is isolated by the above screening using a binding activity of the protein of the invention can also be used as a sample.

[0095] A compound isolated by the screening may be a candidate to be an agonist or an antagonist of the receptor of the protein. By utilizing an assay that monitors a change in the intracellular signaling such as phosphorylation which results from reduction of the binding between the protein and its receptor, it is possible to identify whether the obtained compound is an agonist or antagonist of the receptor. Also, the compound may be a candidate of a molecule that can

inhibit the interaction between the protein and its associated proteins (including a receptor) in vivo. Such compounds can be used for developing drugs for precaution or cures of a disease with which the protein is associated.

[0096] Secretory proteins may regulate cellular conditions such as growth and differentiation. It is possible to find out a novel factor that regulates cellular conditions by adding the secretory protein of the invention to a certain kind of cell, and performing a screening by utilizing the cellular changes in growth or differentiation, or activation of a particular gene.

[0097] The screening can be performed, for example, as follows. First, the protein of the invention is expressed and purified in a recombinant form. Then, the purified protein is microinjected into a various kind of cell lines or primary cultured cells, and the change in the cell growth and differentiation is monitored. The induction of a particular gene that is known to be involved in a certain cellular change is detected with the amounts of mRNA and protein. Alternatively, the amount of an intracellular molecule (low molecular compounds, etc.) that is changed by the function of a gene product (protein) that is known to be functioning in a certain cellular change is used for the detection.

[0098] Once the screening reveals that the protein of the invention can regulate cellular conditions or the functions, it is possible to apply the protein as a pharmaceutical and diagnostic medicine for associated diseases by itself or by altering a part of it into an appropriate composition.

[0099] As is above described for membrane proteins, the secretory protein provided by the invention may be used to explore a novel ligand-receptor interaction using a screening based on the binding activity to a known ligand or receptor. A similar method can be used to identify an agonist or antagonist. The resulting compounds obtained by the methods can be a candidate of a compound that can inhibit the interaction between the protein of the invention and an interacting molecule (including a receptor). The compounds may be able to use as a preventive, therapeutic, and diagnostic medicine for the diseases, in which the protein may play a certain role.

[0100] Proteins associated with signal transduction or transcription may be a factor that affects a certain protein or gene in response to intracellular/extracellular stimuli. It is possible to find out a novel factor that can affect a protein or gene by expressing the protein provided by the invention in a certain types of cells, and performing a screening utilizing the activation of a certain intracellular protein or gene.

[0101] The screening may be performed as follows. First, a transformed cell expressing the protein is obtained. Then, the transformed cell line and the untransformed original cell are compared for the changes in the expression of a certain gene by detecting the amount of its mRNA or protein. Alternatively, the amount of an intracellular molecule (low molecular compounds), which is changed by the function of a gene product (protein) that is known to function in a certain cellular change, may be used for the detection. Furthermore, the change of the expression of a certain gene can be detected by introducing a fusion gene that comprises a regulatory region of the gene and a marker gene (luciferase, beta-galactosidase, etc.) into a cell, expressing the protein provided by the invention into the cell, and estimating the activity of a marker gene product (protein).

[0102] If the protein or gene of the invention is associated with diseases, it is possible to screen a gene or compound that can regulate its expression and/or activity either directly or indirectly by utilizing the protein of the present invention.

[0103] For example, the protein of the invention is expressed and the recombinant protein is purified. Then, the protein and gene whose expression was affected in the presence of the protein of the invention is also purified, and the binding activity between the two proteins or genes is examined. The examination may be performed with pretreatment with a compound that is candidate of an inhibitor. In an alternative method, a transcription regulatory region locating in the 5'-upstream of the gene encoding the protein of the invention that is capable of regulating the expression of other genes is obtained, and fused with a marker gene. The fusion is introduced into a cell, and the cell is added with compounds to explore a regulatory factor of the expression of the said gene.

[0104] The compound obtained by the screening can be used for developing pharmaceutical and diagnostic medicines for the diseases with which the protein of the present invention is associated. Similarly, if the regulatory factor obtained by the screening is a protein, the protein itself can be used as a pharmaceutical, and if there is a compound that affects the original expression level and/or activity of the protein, it also can be used for the same purpose.

[0105] If the protein of the invention has an enzymatic activity, regardless of whether it is a secretory protein, membrane protein, or proteins associated with signal transduction, glycoprotein, transcription, or diseases, a screening may be performed by adding a compound to the protein of the invention under the suitable condition and monitoring the change of the compound. The enzymatic activity may also be utilized to screen for a compound that can inhibit the activity of the protein.

[0106] In a screening given as an example, the protein of the invention is expressed and the recombinant protein is purified. Then, compounds are contacted with the purified protein, and the amount of the compound and the reaction products is examined. Alternatively, compounds that are candidates of an inhibitor are pretreated, then a compound (substrate) that can react with the purified protein is added, and the amount of the substrate and the reaction products is examined.

[0107] The compounds obtained in the screening may be used as a medicine for diseases with which the protein of the invention is associated. Also they can be applied for tests that examine whether the protein of the invention functions

normally *in vivo*.

[0108] Whether the secretory or membrane protein of the present invention is a novel protein associated with diseases or not is determined in another method than described above, by obtaining a specific antibody against the protein of the invention, and examining the relationship between the expression or activity of the protein and a certain disease.

In an alternative way, it may be analyzed referred to the methods in "Molecular Diagnosis of Genetic Diseases" (Elles R. edit, (1996) in the series of "Method in Molecular Biology" (Humana Press).

[0109] Disease-associated proteins are a target of the above described screenings and very useful for developing a drug that is capable of regulating the expression and activity of the protein. Also, they are useful in medicinal industry as a diagnostic marker of the related disease and as a target for gene therapy.

[0110] Compounds isolated as mentioned above can be administered patients as it is, or after formulated into a pharmaceutical composition according to the known methods. For example, a pharmaceutically acceptable carrier or vehicle, specifically sterilized water, saline, plant oil, emulsifier, or suspending agent can be mixed with the compounds appropriately. The pharmaceutical compositions can be administered to patients by a method known to those skilled in the art, such as intraarterial intravenous, or subcutaneous injections. The dosage may vary depending on the weight or age of a patient, or the method of administration, but those skilled in the art can choose an appropriate dosage properly. If the compound is encoded by DNA, the DNA can be cloned into a vector for gene therapy, and used for gene therapy. The dosage of the DNA and the method of its administration may vary depending on the weight or age of a patient, or the symptoms, but those skilled in the art can choose properly.

[0111] The protein encoded by the polynucleotide of the invention can be prepared as a recombinant protein or as a natural protein. For example, the recombinant protein can be prepared by inserting the polynucleotide encoding the protein of the invention into a vector, introducing the vector into an appropriate host cell and purifying the protein expressed within the transformed host cell, as described below. In contrast, the natural protein can be prepared, for example, by utilizing an affinity column to which an antibody against the protein of the invention (Current Protocols in Molecular Biology (1987) Ausubel et al. edit, John Wiley & Sons, Section 16.1-16.19) is attached. The antibody used for the affinity purification may be either a polyclonal antibody, or a monoclonal antibody. Alternatively, *in vitro* translation (See, for example, "On the fidelity of mRNA translation in the nuclease-treated rabbit reticulocyte lysate system." Dasso M.C., and Jackson R.J. (1989) Nucleic Acids Res. 17: 3129-3144) may be used for preparing the protein of the invention.

[0112] Proteins functionally equivalent to the proteins of the present invention can be prepared based on the activities, which were clarified in the above-mentioned manner, of the proteins of the present invention. Using the biological activity possessed by the protein of the invention as an index, it is possible to verify whether or not a particular protein is functionally equivalent to the protein of the invention by examining whether or not the protein has said activity.

[0113] Proteins functionally equivalent to the proteins of the present invention can be prepared by those skilled in the art, for example, by using a method for introducing mutations into an amino acid sequence of a protein (for example, site-directed mutagenesis (Current Protocols in Molecular Biology, edit, Ausubel et al., (1987) John Wiley & Sons, Section 8.1-8.5). Besides, such proteins can be generated by spontaneous mutations. The present invention comprises the proteins having one or more amino acid substitutions, deletions, insertions and/or additions in the amino acid sequences of the proteins of the present invention (Table 370), as far as the proteins have the equivalent functions to those of the proteins identified in the present Examples described later.

[0114] There are no limitations in the number and sites of amino acid mutations, as far as the proteins maintain the functions thereof. The number of mutations is typically 30% or less, or 20% or less, or 10% or less, preferably within 5% or less, or 3% or less of the total amino acids, more preferably within 2% or less or 1% or less of the total amino acids. From the viewpoint of maintaining the protein function, it is preferable that a substituted amino has a similar property to that of the original amino acid. For example, Ala, Val, Leu, Ile, Pro, Met, Phe and Trp are assumed to have similar properties to one another because they are all classified into a group of non-polar amino acids. Similarly, substitution can be performed among non-charged amino acid such as Gly, Ser, Thr, Cys, Tyr, Asn, and Gln, acidic amino acids such as Asp and Glu, and basic amino acids such as Lys, Arg, and His.

[0115] In addition, proteins functionally equivalent to the proteins of the present invention can be isolated by using techniques of hybridization or gene amplification known to those skilled in the art. Specifically, using the hybridization technique (Current Protocols in Molecular Biology, edit, Ausubel et al., (1987) John Wiley & Sons, Section 6.3-6.4), those skilled in the art can usually isolate a DNA highly homologous to the DNA encoding the protein identified in the present Example based on the identified nucleotide sequence (Table 370) or a portion thereof and obtain the functionally equivalent protein from the isolated DNA. The present invention includes proteins encoded by the DNAs hybridizing with the DNAs encoding the proteins identified in the present Example, as far as the proteins are functionally equivalent to the proteins identified in the present Example. Organisms from which the functionally equivalent proteins are isolated are illustrated by vertebrates such as human, mouse, rat, rabbit, pig and bovine, but are not limited to these animals.

[0116] Washing conditions of hybridization for the isolation of DNAs encoding the functionally equivalent proteins are usually "1xSSC, 0.1% SDS, 37°"; more stringent conditions are "0.5xSSC, 0.1% SDS, 42°"; and still more stringent conditions are "0.1 x SSC, 0.1% SDS, 65°". Alternatively, the following conditions can be given as hybridization con-

ditions of the present invention. Namely, conditions in which the hybridization is done at "6xSSC, 40% Formamide, 25.", and the washing at "1xSSC, 55." can be given. More preferable conditions are those in which the hybridization is done at "6xSSC, 40% Formamide, 37.", and the washing at "0.2xSSC, 55.". Even more preferable are those in which the hybridization is done at "6xSSC, 50% Formamide, 37.", and the washing at "0.1xSSC, 62.". The more stringent the conditions of hybridization are, the more frequently the DNAs highly homologous to the probe sequence are isolated. Therefore, it is preferable to conduct hybridization under stringent conditions. Examples of stringent conditions in the present invention are, washing conditions of "0.5xSSC, 0.1% SDS, 42.", or alternatively, hybridization conditions of "6xSSC, 40% Formamide, 37.", and the washing at "0.2xSSC, 55.". However, the above-mentioned combinations of SSC, SDS and temperature conditions are indicated just as examples. Those skilled in the art can select the hybridization conditions with similar stringency to those mentioned above by properly combining the above-mentioned or other factors (for example, probe concentration, probe length and duration of hybridization reaction) that determines the stringency of hybridization.

[0117] The amino acid sequences of proteins isolated by using the hybridization techniques usually exhibit high homology to those of the proteins of the present invention, which are shown in Table 370. The present invention encompasses a polynucleotide comprising a nucleotide sequence that has a high identity to the nucleotide sequence of claim 8 (a). Furthermore, the present invention encompasses a peptide, or protein comprising an amino acid sequence that has a high identity to the amino acid sequence encoded by the polynucleotide of claim 8 (b). The term "high identity" indicates sequence identity of at least 40% or more; preferably 60% or more; and more preferably 70% or more. Alternatively, more preferable is identity of 90% or more, or 93% or more, or 95% or more, furthermore, 97% or more, or 99% or more. The identity can be determined by using the BLAST search algorithm.

[0118] With the gene amplification technique (PCR) (Current Protocols in Molecular Biology, edit, Ausubel et al., (1987) John Wiley & Sons, Section 6.3-6.4)) using primers designed based on the nucleotide sequence (Table 370) or a portion thereof identified in the present Example, it is possible to isolate a DNA fragment highly homologous to the nucleotide sequence or a portion thereof and to obtain functionally equivalent protein to a particular protein identified in the present Example based on the isolated DNA fragment.

[0119] The "percent identity" of two amino acid sequences or of two nucleic acids is determined using the algorithm of Karlin and Altschul (Proc. Natl. Acad. Sci. USA 87:2264-2268, 1990), modified as in Karlin and Altschul (Proc. Natl. Acad. Sci. USA 90:5873-5877, 1993). Such an algorithm is incorporated into the BLASTN and BLASTX programs of Altschul et al. (J. Mol. Biol. 215:403-410, 1990). BLAST nucleotide searches are performed with the BLASTN program, score = 100, wordlength = 12. BLAST protein searches are performed with the BLASTX program, score = 50, wordlength = 3. When gaps exist between two sequences, Gapped BLAST is utilized as described in Altschul et al. (Nucleic Acids Res. 25:3389-3402, 1997). When utilizing BLAST and Gapped BLAST programs, the default parameters of the respective programs (e.g., BLASTX and BLASTN) are used. See <http://www.ncbi.nlm.nih.gov>.

[0120] The present invention also includes a partial peptide of the proteins of the invention. The partial peptide comprises a protein generated as a result that a signal peptide has been removed from a secretory protein. If the protein of the present invention has an activity as a receptor or a ligand, the partial peptide may function as a competitive inhibitor of the protein and may bind to the receptor (or ligand). In addition, the present invention comprises an antigen peptide for raising antibodies. For the peptides to be specific for the protein of the invention, the peptides comprise at least 7 amino acids, preferably 8 amino acids or more, more preferably 9 amino acids or more, and even more preferably 10 amino acids or more. The peptide can be used for preparing antibodies against the protein of the invention, or competitive inhibitors of them, and also screening for a receptor that binds to the protein of the invention. The partial peptides of the invention can be produced, for example, by genetic engineering methods, known methods for synthesizing peptides, or digesting the protein of the invention with an appropriate peptidase.

[0121] The present invention also relates to a vector into which the DNA of the invention is inserted. The vector of the invention is not limited as long as it contains the inserted DNA stably. For example, if *E. coli* is used as a host, vectors such as pBluescript vector (Stratagene) are preferable as a cloning vector. To produce the protein of the invention, expression vectors are especially useful. Any expression vector can be used as far as it is capable of expressing the protein in vitro, in *E. coli*, in cultured cells, or in vivo. For example, pBEST vector (Promega) is preferable for in vitro expression, pET vector (Invitrogen) for *E. coli*, pME18S-FL3 vector (GenBank Accession No. AB009864) for cultured cells, and pME18S vector (Mol. Cell. Biol. (1988) 8: 466-472) for in vivo expression. To insert the DNA of the invention, ligation utilizing restriction sites can be performed according to the standard method (Current Protocols in Molecular Biology (1987) Ausubel et al. edit, John Wiley & Sons, Section 11.4-11.11).

[0122] The present invention also relates to a transformant carrying the vector of the invention. Any cell can be used as a host into which the vector of the invention is inserted, and various kinds of host cells can be used depending on the purposes. For strong expression of the protein in eukaryotic cells, COS cells or CHO cells can be used, for example.

[0123] Introduction of the vector into host cells can be performed, for example, by calcium phosphate precipitation method, electroporation method (Current Protocols in Molecular Biology (1987) Ausubel et al. edit, John Wiley & Sons, Section 9.1-9.9), lipofectamine method (GIBCO-BRL), or microinjection method, etc.

[0124] The primer of the present invention can be used for synthesizing full-length cDNA, and also for the detection and/or diagnosis of the abnormality of the protein of the invention encoded by the full-length cDNA. For example, by utilizing polymerase chain reaction (genomic DNA-PCR, or RT-PCR) using the primer of the invention, DNA encoding the protein of the invention can be amplified. It is also possible to obtain the regulatory region of expression in the 5'-upstream by using PCR or hybridization since the transcription start site within the genomic sequence can be easily specified based on the 5'-end sequence of the full-length cDNA. The obtained genomic region can be used for detection and/or diagnosis of the abnormality of the sequence by RFLP analysis, SSCP, or direct sequencing.

[0125] Furthermore, the "polynucleotide having a length of at least 15 nucleotides, comprising a nucleotide sequence that is complementary to a polynucleotide comprising the nucleotide sequence set forth in any one of SEQ ID NOs in Table 370, or its complementary strand" includes an antisense polynucleotide for suppressing the expression of the protein of the invention. To exert the antisense effect, the antisense polynucleotide has a length of at least 15 bp or more, for example, 50 bp or more, preferably 100 bp or more, and more preferably 500 bp or more, and has a length of usually 3000 bp or less and preferably 2000 bp or less. The antisense DNA can be used in the gene therapy of the diseases that are caused by the abnormality of the protein of the invention (abnormal function or abnormal expression). Said antisense DNA can be prepared, for example, by the phosphorothioate method ("Physicochemical properties of phosphorothioate oligodeoxynucleotides." Stein (1988) Nucleic Acids Res. 16: 3209-3221) based on the nucleotide sequence of the DNA encoding the protein (for example, the DNA set forth in any one of SEQ ID NOs in Table 370).

[0126] The polynucleotide or antisense DNA of the present invention can be used in gene therapy, for example, by administering it into a patient by the in vivo or ex vivo method with virus vectors such as retrovirus vectors, adenovirus vectors, and adeno-associated virus vectors, or non-virus vectors such as liposome.

[0127] The present invention also relates to antibodies that bind to the protein of the invention. There are no limitations in the form of the antibodies of the invention. They include polyclonal antibodies, monoclonal antibodies, or their portions that can bind to the protein of the invention. They also include antibodies of all classes. Furthermore, special antibodies such as humanized antibodies are also included.

[0128] The polyclonal antibody of the invention can be obtained according to the standard method by synthesizing an oligopeptide corresponding to the amino acid sequence and immunizing rabbits with the peptide (Current Protocols in Molecular Biology (1987) Ausubel et al. edit, John Wiley & Sons, Section 11.12-11.13). The monoclonal antibody of the invention can be obtained according to the standard method by purifying the protein expressed in E. coli, immunizing mice with the protein, and producing a hybridoma cell by fusing the spleen cells and myeloma cells (Current Protocols in Molecular Biology (1987) Ausubel et al. edit, John Wiley & Sons, Section 11.4-11.11).

[0129] The antibody binding to the protein of the present invention can be used for purification of the protein of the invention, and also for detection and/or diagnosis of the abnormalities of the expression and structure of the protein. Specifically, proteins can be extracted, for example, from tissues, blood, or cells, and the protein of the invention is detected by Western blotting, immunoprecipitation, or ELISA, etc. for the above purpose.

[0130] Furthermore, the antibody binding to the protein of the present invention can be utilized for treating the diseases that associates with the protein of the invention. If the antibodies are used for treating patients, human antibodies or humanized antibodies are preferable in terms of their low antigenicity. The human antibodies can be prepared by immunizing a mouse whose immune system is replaced with that of human ("Functional transplant of megabase human immunoglobulin loci recapitulates human antibody response in mice" Mendez M.J. et al. (1997) Nat. Genet. 15: 146-156). The humanized antibodies can be prepared by recombination of the hypervariable region of a monoclonal antibody (Methods in Enzymology (1991) 203: 99-121).

[0131] The present invention further relates to databases comprising at least a sequence of polynucleotide and/or protein, or a medium recorded in such databases, selected from the sequence data of the nucleotide and/or the amino acids indicated in Table 370. The term "database" means a set of accumulated information as machine-searchable and readable information of nucleotide sequence. The databases of the present invention comprise at least one of the novel nucleotide sequences of polynucleotide provided by the present invention. The databases of the present invention can consist of only the sequence data of the polynucleotide provided by the present invention or can comprise other information on nucleotide sequences of known full-length cDNAs or ESTs. The databases of the present invention can be comprised of not only the information on the nucleotide sequences but also the information on the gene functions revealed by the present invention. Additional information such as names of DNA clones carrying the full-length cDNAs can be recorded or linked together with the sequence data in the databases.

[0132] The database of the present invention is useful for gaining complete gene sequence information from partial sequence information of a gene of interest. The database of the present invention comprises nucleotide sequence information of full-length cDNAs. Consequently, by comparing the information in this database with the nucleotide sequence of a partial gene fragment yielded by differential display method or subtraction method, the information on the full-length nucleotide sequence of interest can be gained from the sequence of the partial fragment as a starting clue.

[0133] The sequence information of the full-length cDNAs constituting the database of the present invention contains not only the information on the complete sequences but also extra information on expression frequency of the genes

as well as homology of the genes to known genes and known proteins. Thus the extra information facilitates rapid functional analyses of partial gene fragments. Further, the information on human genes is accumulated in the database of the present invention, and therefore, the database is useful for isolating a human homologue of a gene originating from other species. The human homologue can be isolated based on the nucleotide sequence of the gene from the original species.

[0134] At present, information on a wide variety of gene fragments can be obtained by differential display method and subtraction method. In general, these gene fragments are utilized as tools for isolating the full-length sequences thereof. When the gene fragment corresponds to an already-known gene, the full-length sequence is easily obtained by comparing the partial sequence with the information in known databases. However, when there exists no information corresponding to the partial sequence of interest in the known databases, cDNA cloning should be carried out for the full-length cDNA. It is often difficult to obtain the full-length nucleotide sequence using the partial sequence information as an initial clue. If the full-length of the gene is not available, the amino acid sequence of the protein encoded by the gene remains unidentified. Thus the database of the present invention can contribute to the identification of full-length cDNAs corresponding to gene fragments, which cannot be revealed by using databases of known genes.

[0135] The invention is illustrated more specifically with reference to the following examples, but is not to be construed as being limited thereto.

EXAMPLE 1

Construction of a cDNA library by the oligo-capping method.

[0136] The NT-2 neuron progenitor cells (Stratagene), a teratocarcinoma cell line from human embryo testis, which can differentiate into neurons by treatment with retinoic acid were used. The NT-2 cells were cultured according to the manufacturer's instructions as follows.

- (1) NT-2 cells were cultured without induction by retinoic acid treatment ((NT2RM1, NT2RM2, NT2RM4)).
- (2) After cultured, NT-2 cells were induced by adding retinoic acid, and then were cultured for 48 hours (NT2RP1).
- (3) After cultured, NT-2 cells were induced by adding retinoic acid, and then were cultured for 2 weeks (NT2RP2, NT2RP3, NT2RP4).

[0137] Also, the human brain neuroglioma cell line H4 (ATCC HTG-148) (BNGH41), human neuroblastoma cell line SK-N-MC (ATCC HTB-10) (SKNMC1), and human retinoblastoma cell line Y79 (ATCC HTB-18) (Y79AA1) were cultured according to the culture conditions described in the ATCC catalogue. The cells were harvested separately, and mRNA was extracted from each cell by the method described in the literature (Molecular Cloning 2nd edition. Sambrook J., Fritsch, E.F., and Maniatis T. (1989) Cold Spring Harbor Laboratory Press). Furthermore, poly(A)⁺RNA was purified from the mRNA using oligo-dT cellulose.

[0138] Similarly, human placenta (PLACE1, PLACE2, PLACE3), human ovary cancer tissue (OVARC1), tissues from human embryo at 10 weeks, which is enriched with head (HEMBA1), or body (HEMBB1), human mammary gland (MAMMA1), human thyroid gland (THYRO1) were used to extract mRNA by the method described in the literature (Molecular Cloning 2nd edition. Sambrook J., Fritsch, E.F., and Maniatis T. (1989) Cold Spring Harbor Laboratory Press). Furthermore, poly(A)⁺RNA was purified from the mRNA using oligo-dT cellulose.

[0139] Each poly(A)⁺RNA was used to construct a cDNA library by the oligo-capping method (Maruyama M. and Sugano S. (1994) Gene 138: 171-174). Using the Oligo-cap linker (SEQ ID NO: 2541) and the Oligo-dT primer (SEQ ID NO: 2542), bacterial alkaline phosphatase (BAP) treatment, tobacco acid phosphatase (TAP) treatment, RNA ligation, the first strand cDNA synthesis, and removal of RNA were performed as described in the reference (Suzuki and Kanno (1996) Protein Nucleic acid and Enzyme. 41: 197-201; Suzuki Y. et al. (1997) Gene 200: 149-156). Next, 5'- and 3'-PCR primers (SEQ ID NO: 2543, and 2544, respectively) were used for performing PCR to convert the cDNA into double stranded cDNA, which was then digested with SfiI. Then, the DraIII-cleaved pUC19FL3 vector (Figure 1; for NT2RM1, and NT2RP1), or the DraIII-cleaved pME18SFL3 (Figure 1) (GenBank AB009864, expression vector; for NT2RM2, NT2RM4, NT2RP2, NT2RP3, NT2RP4, BNGH41, SKNMC1, Y79AA1, PLACE1, PLACE2, PLACE3, OVARC1, HEMBA1, HEMBB1, MAMMA1, and THYRO1) was used for cloning the cDNA in an unidirectional manner, and cDNA libraries were obtained. Then, the nucleotide sequence of the 5'- and 3'- ends of the cDNA clones was analyzed with a DNA sequencer (ABI PRISM 377, PE Biosystems) after sequencing reactions were performed with the DNA sequencing reagents (Dye Terminator Cycle Sequencing FS Ready Reaction Kit, dRhodamine Terminator Cycle Sequencing FS Ready Reaction Kit, or BigDye Terminator Cycle Sequencing FS Ready Reaction Kit, from PE Biosystems) according to the instructions. The data were compiled into a database.

[0140] The full-length-enriched cDNA libraries except those for NT2RM1 and NT2RP1 were constructed using eukaryotic expression vector pME18SFL3. The vector contains SR α promoter and SV40 small t intron in the upstream

of the cloning site, and SV40 polyA added signal sequence site in the downstream. As the cloning site of pME18SFL3 has asymmetrical DraIII sites, and the ends of cDNA fragments contain SfiI sites complementary to the DraIII sites, the cloned cDNA fragments can be inserted into the downstream of the SR α promoter unidirectionally. Therefore, clones containing full-length cDNA can be expressed transiently by introducing the obtained plasmid directly into COS cells. Thus, the clones can be analyzed very easily in terms of the proteins that are the gene products of the clones, or in terms of the biological

[0141] Herein, the cDNA libraries and the name of each clone are related as shown in Table 2. Therein, "xxxxxx" represents the clone number of six digits. Thus, the sequences are named by the library name, the clone number plus F- for the 5'-end, or R- for the 3'-end.

Table 2

library:

clone	5'-end sequence	3'-end sequence
NT2RM1:		
NT2RM1xxxxxx	F-NT2RM1xxxxxx	
NT2RP1:		
NT2RP1xxxxxx	F-NT2RP1xxxxxx	
NT2RM2:		
NT2RM2xxxxxx	F-NT2RM2xxxxxx	R-NT2RM2xxxxxx
NT2RM4:		
NT2RM4xxxxxx	F-NT2RM4xxxxxx	R-NT2RM4xxxxxx
NT2RP2:		
NT2RP2xxxxxx	F-NT2RP2xxxxxx	R-NT2RP2xxxxxx
NT2RP3:		
NT2RP3xxxxxx	F-NT2RP3xxxxxx	R-NT2RP3xxxxxx
NT2RP4:		
NT2RP4xxxxxx	F-NT2RP4xxxxxx	R-NT2RP4xxxxxx
BNGH41:		
BNGH41xxxxxx	F-BNGH41xxxxxx	R-BNGH41xxxxxx
SKNMC1:		
SKNMC1xxxxxx	F-SKNMC1xxxxxx	R-SKNMC1xxxxxx
Y79AA1:		
Y79AA1xxxxxx	F-Y79AA1xxxxxx	R-Y79AA1xxxxxx
PLACE1:		
PLACE1xxxxxx	F-PLACE1xxxxxx	R-PLACE1xxxxxx
PLACE2:		
PLACE2xxxxxx	F-PLACE2xxxxxx	R-PLACE2xxxxxx
PLACE3:		
PLACE3xxxxxx	F-PLACE3xxxxxx	R-PLACE3xxxxxx
OVARC1:		
OVARC1xxxxxx	F-OVARC1xxxxxx	R-OVARC1xxxxxx
HEMBA1:		
HEMBA1xxxxxx	F-HEMBA1xxxxxx	R-HEMBA1xxxxxx
HEMBB1:		
HEMBB1xxxxxx	F-HEMBB1xxxxxx	R-HEMBB1xxxxxx
MAMMA1:		
MAMMA1xxxxxx	F-MAMMA1xxxxxx	R-MAMMA1xxxxxx
THYRO1:		

THYRO1xxxxxx F-THYRO1xxxxxx R-THYRO1xxxxxx

EXAMPLE 2

Estimation of the fullness ratio of the 5'-ends of the clones contained in the cDNA libraries constructed by the oligo-capping method.

[0142] The fullness ratio at the 5'-end sequences of the 59,823 clones in the human cDNA libraries constructed by the oligo-capping method was determined as follows. Of all the clones whose 5'-end sequences were found in those of known human mRNA in the public database, a clone was judged to be "full-length", if it had a longer 5'-end sequence than that of the known human mRNA, or, even though the 5'-end sequence was shorter, if it contained the translation initiation codon. A clone which did not contain the translation initiation codon was judged to be "non-full-length". The fullness ratio ((the number of full-length clones)/(the number of full-length and non-full-length clones)) at the 5'-end of the cDNA clones from each library was determined by comparing with the known human mRNA. As a result, the fullness ratio of the 5'-ends was 63.5%. It suggests that the human cDNA clones obtained by the described method have complete 5'-ends with high probability.

EXAMPLE 3

Assessment of the fullness ratio of the 5'-end of the cDNA by the ATGpr and the ESTiMateFL.

[0143] The ATGpr, developed by Salamov A.A., Nishikawa T., and Swindells M.B. in the Helix Research Institute, is a program for prediction of the translation start codon based on the characteristics of the sequences in the vicinity of the ATG codon. The results are shown with expectations that an ATG is a true start codon (0.05-0.94). When this program is applied to general cDNAs without considering whether or not the ATG codons in the cDNAs are the true initiation codons of the cDNAs, both the sensitivity and the specificity of the results are estimated at 66%. Here, the sensitivity means the ratio of the number of codons judged to be initiation codons by the program to the total number of true initiation codons, and the specificity means the ratio of the number of true initiation codons to the number of codons judged to be initiation codons by the program. In contrast, when the program was applied to the 5'-end sequences of the clones from the cDNA library that was obtained by the oligo-capping method and that had 65% fullness ratio, the sensitivity and specificity of evaluation of a full-length clone (clone containing the N-terminal end of ORF) were improved to 82-83% by selecting only clones having the ATGpr1 score 0.6 or higher.

[0144] Furthermore, the program was used to assess the fullness of 18,959 clones in the human cDNA libraries obtained here, which have 5'-ends matched to a known human mRNA. Briefly, the maximal ATGpr1 score of the clones was determined, and then their 5'-end sequence was compared with the known human mRNA to estimate whether the clone is full-length or not. The result was summarized in Table 3. Based on the knowledge that known mRNAs, in general, are highly expressed in the cell, the expression levels of genes having a low number in the EST hit, which represent mRNAs whose expression levels are relatively low were examined, and the result is shown in Table 4.

[0145] In the table, the number of full-length clones indicate that of clones containing the N-terminal end of ORF, and so does the number of non-full-length clones that of clones without the N-terminal end of ORF. The fullness ratio represents (the number of full-length clones)/(the number of full-length clones plus the number of non-full-length clones).

Table 3

The maximal ATGpr1 score and the fullness ratio of the 5'-end sequences of clones obtained from human cDNA libraries constructed by the oligo-capping method; clones having a matched 5'-end with that of a known human mRNA.			
maximal ATGpr1 score	number of (full-length clones plus non-full-length clones)	number of full-length clones	fullness ratio
>=0.70	11,193	9,346	83.5%
>=0.50	13,369	10,549	78.9%
>=0.30	15,489	11,340	73.2%

EP 1 130 094 A2

Table 3 (continued)

5	The maximal ATGpr1 score and the fullness ratio of the 5'-end sequences of clones obtained from human cDNA libraries constructed by the oligo-capping method; clones having a matched 5'-end with that of a known human mRNA.			
	maximal ATGpr1 score	number of (full-length clones plus non-full-length clones)	number of full-length clones	fullness ratio
10	>=0.15	17,394	11,811	67.9%
	>=0.00	18,959	12,046	63.5%

Table 4

15	The maximal ATGpr1 score and the fullness ratio of the 5'-end sequences of the clones obtained from human cDNA libraries constructed by the oligo-capping method; clones having 5 EST hits or less among the clones having a matched 5'-end with that of a known human mRNA.			
	maximal ATGpr1 score	number of (full-length clones plus non-full-length clones)	number of full-length clones	fullness ratio
20	>=0.70	2,801	1,934	69.0%
	>=0.50	3,683	2,393	65.0%
25	>=0.30	4,683	2,707	57.8%
	>=0.15	5,559	2,890	52.0%
	>=0.00	6,113	3,013	49.8%

[0146] The ESTiMateFL, developed by Nishikawa and Ota in the Helix Research Institute, is a method for the selection of a clone with high fullness ratio by comparing with the 5'-end or 3'-end sequences of ESTs in the public database.

[0147] By the method, a cDNA clone is judged presumably not to be full-length if there exist any ESTs which have longer 5'-end or 3'-end sequences than the clone. The method is systematized for high throughput analysis. A clone is judged to be full-length if the clone has a longer 5'-end sequence than ESTs in the public database. Even if a clone has a shorter 5'-end, the clone is judged to be full-length if the difference in length is within 50 bases, and otherwise judged not to be full-length, for convenience. In case of the 5'-end sequence of the clones which matches a known mRNA, about 80% of the sequences that were judged to be full-length by comparing with ESTs was judged to be full-length by estimating the 5'-end sequence, as well; about 80% of the sequences that were judged to be not full-length by comparing with ESTs was judged to be not full-length by estimating the 5'-end sequence, as well. The accuracy of the prediction by comparing cDNA clones with ESTs is improved with increasing number of ESTs to be compared. However, when only a limited number of ESTs are available, the reliability becomes low. Thus, the method is effective in excluding clones with high probability of being non-full-length, from the cDNA clones that is synthesized by the oligo-capping method and that have the 5'-end sequences with about 60 % fullness ratio. In particular, the ESTiMateFL is efficiently used to estimate the fullness ratio at the 3'-end sequence of cDNA of a human unknown mRNA which has a significant number of ESTs in the public database.

[0148] The 18,959 clones isolated from human cDNA libraries constructed by the oligo-capping method, which have the 5'-end sequence that matches a known human mRNA, were estimated by using the ATGpr and ESTiMateFL. Briefly, the 5'-end sequence that matches a known human mRNA of the respective clone was analyzed to obtain the maximal ATGpr1 score, and compared with the ORF of the known human mRNA that matches it to determine whether the clone is full-length or not. Then, the 5'-end sequence of the respective clone was analyzed by the ESTiMateFL to judge whether the clone is full-length or not. Specifically, the 5'-end sequences that match a known human mRNA of the 18,959 clones constructed by the oligo-capping method were compared with those of ESTs by the ESTiMateFL and the clones other than those that are not full-length were selected. Then, the selected clones were used to analyze the relationship between the ATGpr and the fullness ratio. The result was summarized in Table 5. Also, among the selected, the clones in which the number of the EST hit is not more than 5 were selected and analyzed. The result was summarized in Table 6, which represents the result of the analysis of mRNA with relatively low abundance.

[0149] In the Tables, the number of full-length clones, the number of non-full-length clones, and the fullness ratio indicate the number of the clones that contain the N-terminus of the ORF, the number of the clones that do not contain

EP 1 130 094 A2

the N-terminus of the ORF, and (the number of full-length clones)/(the number of full-length clones) plus (the number of non-full-length clones), respectively.

Table 5

The maximal ATGpr1 score and the fullness ratio of the 5'-end sequence in the clones isolated from human cDNA libraries constructed by the oligo-capping method, which have the 5'-end sequence that matches a known human mRNA, and also other than those being not full-length according to the comparison with ESTs.			
maximal ATGpr1 score	number of (full-length clones plus non-full-length clones)	number of full-length clones	fullness ratio
>=0.70	9,068	8,349	92.1%
>=0.50	10,345	9,318	90.1%
>=0.30	11,425	9,964	87.2%
>=0.15	12,254	10,335	84.3%
>=0.00	12,785	10,484	82.0%

Table 6

The maximal ATGpr1 score and the fullness ratio of the 5'-end sequence in the clones isolated from human cDNA libraries constructed by the oligo-capping method, which have the 5'-end sequence that matches a known human mRNA, and also other than those being not full-length according to the comparison with ESTs, in which the number of the EST hit is not more than 5.			
maximal ATGpr1 score	number of (full-length clones plus non-full-length clones)	number of full-length clones	fullness ratio
>=0.70	1,959	1,510	77.1%
>=0.50	2,469	1,821	73.8%
>=0.30	2,975	2,046	68.8%
>=0.15	3,368	2,164	64.3%
>=0.00	3,661	2,226	60.8%

[0150] According to the above results, it was found that, in case of using clones isolated from human cDNA libraries constructed by the oligo-capping method, the fullness ratio of the clones that have low score in the ATGpr can be improved by assessing their 5'-end sequence using the combination of the ATGpr and the ESTiMateFL. Therefore, the method was applied to select a cDNA clone with high fullness ratio.

EXAMPLE 4

Clustering of the 5'-end and 3'-end sequences of cDNA clones.

[0151] The 5'-end and 3'-end sequences of cDNA clones were obtained, and clustered separately. Briefly, data of the single pass sequencing of the determined 5'-end and 3'-end of cDNA clones was subjected to the BLAST search between the sequence data of all the clones synthesized in Example 1, and clones that are supposed to be originating from the same gene were clustered into a group. For the 5'-end sequence, those having the consensus sequence of 95% identity 300 base pairs or more are clustered into the same group. For the 3'-end sequence, those having the consensus sequence of 90% identity 200 base pairs or more are clustered into the same group. Among the clusters of the 5'-end and 3'-end sequences, the sequence having the longest lead was chosen as the representative sequence of the cluster (group).

EXAMPLE 5

Characterization of the representative sequences and the sequences of clones

[0152] Data of the 5'-end sequences of the representative sequences and clones was characterized by the following

EP 1 130 094 A2

methods:

- (1) judging whether it is identical to the sequence of mRNA or ESTs from human by the BLAST search of the GenBank or SwissProt, and examining whether it is full-length by comparing with the sequences of known mRNA and ESTs from human.
- (2) determining the ATGpr1 score using all the initiation codons contained within the 5'-end sequence by the ATGpr which predict fullness ratio.
- (3) predicting the existence of the signal sequence using all the initiation codons contained within the 5'-end sequence by the PSORT which predict signal.
- and,
- (4) only with the 5'-end sequences of the representative sequences of the clusters, examining the keywords in the top hit data of the homology search of the SwissProt.

[0153] Data of the characterized representative sequences and clones was used for the final selection of the clones.

EXAMPLE 6

Identity to the human mRNA and human EST, and comparison of the 5'-end length.

- [0154] The clones and the representative sequences of the clusters were judged to be identical to any human mRNA, if their 5'-end sequence has a region of 200 nucleotides or longer with 94% or more identity to the mRNA. The clones and the representative sequences of the clusters were judged to be identical to any human EST, if their 5'-end sequence has a region of 200 nucleotides or longer with 90% or more identity to the EST.

- [0155] The clones and the representative sequences of the clusters were judged to be full-length in comparison with human mRNA, if their 5'-end sequence is longer than those of the mRNA, or it contains the translation initiation site. The clones and the representative sequences of the clusters were judged to be full-length in comparison with human EST in the database, if their 5'-end sequence is longer than those of the EST, or even though it is shorter, the difference in length between the two sequences is 50 nucleotides or less, for convenience. Otherwise, the clones and the representative sequences of the clusters were judged to be not full-length.

EXAMPLE 7

Prediction of the fullness ratio by the ATGpr.

- [0156] The score in the ATGpr1 is the expectation to be full-length based on calculations, and the higher score reflects the higher fullness ratio as shown in Example 3. Further, the maximal ATGpr1 score represents the score obtained with all the initiation codons contained in the 5'-end sequence of the clones and the representative sequences, and are used for the characterization.

EXAMPLE 8

Prediction of the existence of a signal sequence by the PSORT.

- [0157] Prediction of the existence of a signal sequence by the PSORT was performed on all of the amino acid sequences predicted from all the initiation codons in the 5'-end sequence of the clones and the representative sequences of the clusters. By analyzing the presence or absence of the sequence which is predicted to be a signal sequence, which is characteristics of the N-terminus of many secretory proteins, cDNA clones encoding a secretory protein or membrane protein were selected.

EXAMPLE 9

Prediction of the protein function by the BLAST search.

- [0158] The 5'-end sequence of the representative sequences of the cluster was analyzed by the BLAST homology search of the SwissProt. The obtained top hit data was classified into those identical to the 5'-end representative sequence (identity was 90% or higher), those not identical to the 5'-end representative sequence (identity was 60% or lower, and compared sequence was not more than 25 nucleotides), and those similar to the 5'-end representative sequence (the rest of the data).

EP 1 130 094 A2

[0159] All the keywords in the SwissProt data corresponding to the top hit data were selected, and the 5'-end representative sequences were classified by the keywords relating with functions.

The keywords relating with a secretory protein or membrane protein are the followings:

5 growth factor,
cytokine,
hormone,
receptor,
G-protein coupled receptor,
10 ionic channel,
voltage-gated channel,
calcium channel,
extracellular matrix,
transmembrane, and
15 signal.

[0160] The keywords relating to glycoprotein is glycoprotein.

[0161] The keywords relating to signal transduction are the followings:

20 serine/threonine-protein kinase,
tyrosine-protein kinase, and
calmodulin-binding.

[0162] The keywords relating to transcription are the followings:

25 transcription regulation and activator,
transcription regulation and repressor, and
nuclear protein and repressor.

30 **[0163]** The keywords relating to diseases are disease mutation, and syndrome.

[0164] Many keywords overlapped in the respective group (receptor and transmembrane, for example), and some keywords overlapped in different groups (secretory or membrane, and diseases, etc.).

EXAMPLE 10

35 Selection of clones by characterization.

[0165] From the data obtained by the above characterization, clones encoding a novel secretory protein or membrane protein, or proteins with other predicted functions were selected by the combination of the ATGpr1 score and the prediction of the signal sequence by the PSORT, or according to the top hit data in the homology search of the SwissProt.

40 **[0166]** In selecting the clones, the 5'-end sequences that are identical to a human mRNA were ignored, whereas those that are identical to a human mRNA in part but obviously not identical in the other part were included. Because there were clones selected that are identical to a human mRNA in part but obviously not identical in the other part.

45 **[0167]** Also, if the finally selected clones were found to be not full-length compared with the sequences of human mRNA and ESTs, these clones were discarded.

EXAMPLE 11

50 A method for selection of clones by the combination of the ATGpr1 score and the prediction of the signal sequence by the PSORT (a method for selection of secretory proteins and membrane proteins that are novel and full-length).

[0168] The sequences of clones and the representative sequences of their clusters were used to obtain the maximal ATGpr1 score and predict the presence of the signal sequence. First, clones were selected based on the representative sequences of the clusters. The correspondence between the name and SEQ ID of the representative sequences used for selection (Table 368), and the correspondence between the name and SEQ ID of the introns (including the representative sequences of the 5'-end and 3'-end, and ESTs) used for selection of clones from the representative sequences of the groups (Table 369) were shown in the last part of the present specification. Therein, HRIFA and HRIRA indicate the representative sequence of the 5'-end group, and that of the 3'-end group, respectively.

[0169] In the clusters in which a single clone is contained (the sequence of the 5'-end clone = the representative sequence of the 5'-end), selected were the clones that were judged to be full-length in comparison with human mRNA and ESTs, having the maximal ATGpr1 score 0.5 or higher, and predicted to contain the signal sequence, in principle. However, in the following cases, a clone having a longer 5'-end was selected: the maximal ATGpr1 score was less than 0.5, the sequence of the 5'-end was not full-length, the clone was obviously shorter although the clone was not classified into the same cluster according to the BLAST search of the other clones, or the 5'-end sequence corresponding to the 3'-end of the other clones in the same cluster in which the 3'-end sequence of the clone was contained was found to be longer by assembling. Furthermore, if there were multiple full-length clones in the same cluster and it was not successful to determine by assembling which has the longer 5'-end, all the clones were selected. For assembling, the Sequencher^(TM) (Hitachi Soft Engineering) was used. As a result, the signal sequence predicted to be present in the representative sequence was not present in some of the selected clones. In some cases, the ATGpr1 score became smaller than 0.5 or 0.3. The fullness ratio in these clones was low, yet still it is possible that the clones are full-length. The clones in which the signal sequence predicted to be present in the representative sequence was not present after selection were likely to be without the signal sequence, but still it is possible that the clones encode a membrane protein.

[0170] In the clusters comprising multiple clones, in which the representative sequence of the 5'-end was predicted to contain the signal sequence, selected were the clones having the longest 5'-end sequence among the clones which were judged to be full-length compared with human mRNA and ESTs, having the maximal ATGpr1 score for the 5'-end sequence 0.5 or higher, and predicted to contain the signal sequence. However, in the following cases, a clone having a longer 5'-end was selected: the maximal ATGpr1 score was less than 0.5, the sequence of the 5'-end was not full-length, the clone was obviously shorter although the clone was not classified into the same cluster according to the BLAST search of the other clones, or the 5'-end sequence corresponding to the 3'-end of the other clones in the same cluster in which the 3'-end sequence of the clone was contained was found to be longer. Furthermore, if there were multiple full-length clones in the same cluster and it was not successful to determine by assembling which has the longer 5'-end, all the clones were selected. As a result, the signal sequence predicted to be present in the representative sequence was not present in some of the selected clones. In some cases, the ATGpr1 score became smaller than 0.5 or 0.3. The fullness ratio in these clones was low, yet still it is possible that the clones are full-length. The clones in which the signal sequence predicted to be present in the representative sequence was not present after selection were likely to be without the signal sequence at the 5'-end, but still it is possible that the clones encode a membrane protein.

[0171] Next, in the clusters comprising multiple clones, in which the representative sequence of the 5'-end was predicted to have no signal sequence, selected were the clones which were judged to be full-length compared with human mRNA and ESTs, having the maximal ATGpr1 score for the 5'-end sequence 0.5 or higher, and predicted to contain the signal sequence.

[0172] The number of the clones selected by the combination of the ATGpr1 score and the prediction of a signal sequence by the PSORT were 254. The number of the clones having the maximal ATGpr1 score 0.5 or higher, and predicted to contain a signal sequence were 170 (Table 7-10). Among the clones, 164 clones were found to have the representative sequence of the original cluster that fulfills the same conditions. On the other hand, 5 clones were selected from the representative sequences of the 5'-end of the clusters which was predicted to contain a signal sequence while the maximal ATGpr1 score was lower than 0.5. A clone was selected from the representative sequence of the 5'-end of the cluster which was predicted to have no signal sequence.

[0173] The clones that have the maximal ATGpr1 score 0.3 or higher and less than 0.5 and predicted to contain the signal sequence were 35 clones (Table 11), in which 8 clones were found to have the representative sequence of the original cluster that fulfills the same conditions. Twenty-seven clones were selected from the representative sequences of the clusters which have the maximal ATGpr1 score 0.5 or higher and were predicted to have no signal sequence.

[0174] The clones that have the maximal ATGpr1 score less than 0.3 and were predicted to contain a signal sequence were 41 clones (Table 12). The clones that have the maximal ATGpr1 score 0.5 or higher and were predicted to have no signal sequence were 4 clones (Table 13). The clones that have the maximal ATGpr1 score 0.3 or higher and less than 0.5 and were predicted to have no signal sequence were 2 clones (Table 14). The clones that have the maximal ATGpr1 score less than 0.3 and were predicted to contain a signal sequence were 2 clones (Table 15). The representative sequences of the original clusters of all the clones had the maximal ATGpr1 score 0.3 or higher, and were predicted to contain a signal sequence.

[0175] The fullness ratio of the clones having the maximal ATGpr1 score 0.5 or higher, 0.3 or higher, and 0 or higher is expected to be as shown in Table 3, 4, 5, and 6.

Table 7

EP 1 130 094 A2

The 170 clones in which the selected clones have the maximal ATGpr1 score 0.5 or higher, and were predicted to contain a signal sequence by the PSORT

	name of clone	name of sequence	maximal	signal	name of representative	maximal	signal
			ATGpr1 score		sequence	ATGpr1 score	
5	HEMBA1000713	F-HEMBA1000713	0.67	Yes	HRIFA017729a	0.57	Yes
	HEMBA1000962	F-HEMBA1000962	0.69	Yes	HRIFA000899a	0.69	Yes
	HEMBA1001272	F-HEMBA1001272	0.94	Yes	HRIFA001179a	0.94	Yes
	HEMBA1001297	F-HEMBA1001297	0.89	Yes	HRIFA001201a	0.89	Yes
10	HEMBA1002420	F-HEMBA1002420	0.6	Yes	HRIFA002195a	0.6	Yes
	HEMBA1003101	F-HEMBA1003101	0.67	Yes	HRIFA002787a	0.94	Yes
	HEMBA1003399	F-HEMBA1003399	0.94	Yes	HRIFA002985a	0.94	Yes
	HEMBA1003732	F-HEMBA1003732	0.86	Yes	HRIFA003169a	0.86	Yes
	HEMBA1004110	F-HEMBA1004110	0.59	Yes	HRIFA003379a	0.59	Yes
15	HEMBA1005430	F-HEMBA1005430	0.69	Yes	HRIFA020681a	0.69	Yes
	HEMBA1006016	F-HEMBA1006016	0.6	Yes	HRIFA020466a	0.6	Yes
	HEMBA1006171	F-HEMBA1006171	0.62	Yes	HRIFA021399a	0.62	Yes
	HEMBA1006311	F-HEMBA1006311	0.94	Yes	HRIFA021594a	0.94	Yes
	HEMBA1006335	F-HEMBA1006335	0.83	Yes	HRIFA012069a	0.94	Yes
	HEMBA1006357	F-HEMBA1006357	0.67	Yes	HRIFA021448a	0.67	Yes
20	HEMBA1006658	F-HEMBA1006658	0.66	Yes	HRIFA021323a	0.66	Yes
	HEMBA1006707	F-HEMBA1006707	0.66	Yes	HRIFA021499a	0.94	Yes
	HEMBA1006902	F-HEMBA1006902	0.66	Yes	HRIFA021754a	0.94	Yes
	HEMBA1006960	F-HEMBA1006960	0.94	Yes	HRIFA021886a	0.94	Yes
	HEMBA1000276	F-HEMBA1000276	0.94	Yes	HRIFA029577a	0.94	Yes
25	HEMBA1000642	F-HEMBA1000642	0.94	Yes	HRIFA029779a	0.94	Yes
	HEMBA1000905	F-HEMBA1000905	0.94	Yes	HRIFA009764a	0.91	Yes
	HEMBA1001200	F-HEMBA1001200	0.83	Yes	HRIFA030839a	0.81	Yes
	HEMBA1001407	F-HEMBA1001407	0.87	Yes	HRIFA030981a	0.87	Yes
	HEMBA1001530	F-HEMBA1001530	0.6	Yes	HRIFA031062a	0.6	Yes
	HEMBA1001547	F-HEMBA1001547	0.87	Yes	HRIFA031075a	0.87	Yes
30	HEMBA1001978	F-HEMBA1001978	0.7	Yes	HRIFA031350a	0.7	Yes
	HEMBA1002162	F-HEMBA1002162	0.91	Yes	HRIFA031472a	0.91	Yes
	HEMBA1002228	F-HEMBA1002228	0.53	Yes	HRIFA031510a	0.53	Yes
	HEMBA1002245	F-HEMBA1002245	0.94	Yes	HRIFA032984a	0.94	Yes
	HEMBA1002427	F-HEMBA1002427	0.57	Yes	HRIFA005760a	0.94	Yes
	HEMBA1002465	F-HEMBA1002465	0.72	Yes	HRIFA031672a	0.72	Yes
35	HEMBA1002693	F-HEMBA1002693	0.64	Yes	HRIFA031895a	0.64	Yes
	MAMMA1000046	F-MAMMA1000046	0.7	Yes	HRIFA024841a	0.7	Yes
	MAMMA1000102	F-MAMMA1000102	0.79	Yes	HRIFA026151a	0.79	Yes
	MAMMA1000118	F-MAMMA1000118	0.81	Yes	HRIFA026153a	0.81	Yes
	MAMMA1000141	F-MAMMA1000141	0.8	Yes	HRIFA024554a	0.8	Yes
40	MAMMA1000449	F-MAMMA1000449	0.94	Yes	HRIFA026203a	0.94	Yes
	MAMMA1000457	F-MAMMA1000457	0.78	Yes	HRIFA026210a	0.78	Yes
	MAMMA1000652	F-MAMMA1000652	0.94	Yes	HRIFA026346a	0.94	Yes
	MAMMA1000994	F-MAMMA1000994	0.84	Yes	HRIFA026735a	0.84	Yes
	MAMMA1001141	F-MAMMA1001141	0.89	Yes	HRIFA027265a	0.89	Yes
	MAMMA1001310	F-MAMMA1001310	0.74	Yes	HRIFA026899a	0.74	Yes
45	MAMMA1001344	F-MAMMA1001344	0.71	Yes	HRIFA026918a	0.71	Yes
	MAMMA1002070	F-MAMMA1002070	0.6	Yes	HRIFA028371a	0.82	Yes
	MAMMA1002087	F-MAMMA1002087	0.68	Yes	HRIFA027619a	0.68	Yes
	MAMMA1002165	F-MAMMA1002165	0.57	Yes	HRIFA027673a	0.34	Yes
	MAMMA1002205	F-MAMMA1002205	0.74	Yes	HRIFA027701a	0.74	Yes
50	MAMMA1002633	F-MAMMA1002633	0.53	Yes	HRIFA030461a	0.94	Yes
	NT2RM2000241	F-NT2RM2000241	0.94	Yes	HRIFA020965a	0.94	Yes
	NT2RM2000514	F-NT2RM2000514	0.51	Yes	HRIFA022106a	0.51	Yes
	NT2RM2001643	F-NT2RM2001643	0.69	Yes	HRIFA028926a	0.69	Yes
	NT2RM4000115	F-NT2RM4000115	0.56	Yes	HRIFA025792a	0.53	Yes
	NT2RM4000997	F-NT2RM4000997	0.94	Yes	HRIFA029274a	0.94	Yes

1/4 (continued)

Table 8

The 170 clones in which the selected clones have the maximal ATGpr1 score 0.5 or higher, and were predicted to contain the signal sequence by the PSORT

name of clone	name of sequence	maximal ATGpr1 score	signal	name of representative sequence	maximal ATGpr1 score	signal
NT2RM4001321	F-NT2RM4001321	0.74	Yes	HRIFA024533a	0.74	Yes
NT2RM4001325	F-NT2RM4001325	0.94	Yes	HRIFA033349a	0.94	Yes
NT2RM4001768	F-NT2RM4001768	0.73	Yes	HRIFA013668a	0.5	Yes
NT2RP1000448	F-NT2RP1000448	0.62	Yes	HRIFA005356a	0.62	Yes
NT2RP1001563	F-NT2RP1001563	0.52	Yes	HRIFA006018a	0.52	Yes
NT2RP2001915	F-NT2RP2001915	0.94	Yes	HRIFA007541a	0.94	Yes
NT2RP2002015	F-NT2RP2002015	0.94	Yes	HRIFA007619a	0.94	Yes
NT2RP2002063	F-NT2RP2002063	0.87	Yes	HRIFA007659a	0.94	Yes
NT2RP2002304	F-NT2RP2002304	0.87	Yes	HRIFA007829a	0.94	Yes
NT2RP2002674	F-NT2RP2002674	0.8	Yes	HRIFA008099a	0.8	Yes
NT2RP2002721	F-NT2RP2002721	0.56	Yes	HRIFA008131a	0.56	Yes
NT2RP2003383	F-NT2RP2003383	0.67	Yes	HRIFA008606a	0.67	Yes
NT2RP2003593	F-NT2RP2003593	0.73	Yes	HRIFA008252a	0.94	Yes
NT2RP2003599	F-NT2RP2003599	0.58	Yes	HRIFA008753a	0.58	Yes
NT2RP2003655	F-NT2RP2003655	0.78	Yes	HRIFA008784a	0.83	Yes
NT2RP2004179	F-NT2RP2004179	0.83	Yes	HRIFA008827a	0.83	Yes
NT2RP2004495	F-NT2RP2004495	0.58	Yes	HRIFA009372a	0.58	Yes
NT2RP2004524	F-NT2RP2004524	0.73	Yes	HRIFA009392a	0.82	Yes
NT2RP2004556	F-NT2RP2004556	0.81	Yes	HRIFA009414a	0.81	Yes
NT2RP2004837	F-NT2RP2004837	0.94	Yes	HRIFA006216a	0.93	Yes
NT2RP2005027	F-NT2RP2005027	0.92	Yes	HRIFA004145a	0.93	Yes
NT2RP2005463	F-NT2RP2005463	0.93	Yes	HRIFA010034a	0.42	No
NT2RP2005514	F-NT2RP2005514	0.58	Yes	HRIFA010070a	0.58	Yes
NT2RP2005887	F-NT2RP2005887	0.94	Yes	HRIFA010322a	0.94	Yes
NT2RP2006269	F-NT2RP2006269	0.78	Yes	HRIFA025913a	0.57	Yes
NT2RP3000169	F-NT2RP3000169	0.94	Yes	HRIFA022262a	0.94	Yes
NT2RP3000460	F-NT2RP3000460	0.61	Yes	HRIFA022794a	0.61	Yes
NT2RP3000789	F-NT2RP3000789	0.62	Yes	HRIFA023605a	0.62	Yes
NT2RP3000818	F-NT2RP3000818	0.52	Yes	HRIFA023619a	0.52	Yes
NT2RP3001012	F-NT2RP3001012	0.67	Yes	HRIFA023129a	0.22	Yes
NT2RP3001044	F-NT2RP3001044	0.93	Yes	HRIFA007026a	0.73	Yes
NT2RP3001560	F-NT2RP3001560	0.58	Yes	HRIFA030599a	0.92	Yes
NT2RP3001685	F-NT2RP3001685	0.5	Yes	HRIFA023521a	0.5	Yes
NT2RP3001858	F-NT2RP3001858	0.94	Yes	HRIFA026490a	0.94	Yes
NT2RP3002160	F-NT2RP3002160	0.61	Yes	HRIFA005760a	0.94	Yes
NT2RP3002836	F-NT2RP3002836	0.68	Yes	HRIFA024392a	0.72	Yes
NT2RP3002958	F-NT2RP3002958	0.54	Yes	HRIFA017670a	0.91	Yes
NT2RP3003535	F-NT2RP3003535	0.94	Yes	HRIFA025498a	0.94	Yes
NT2RP3004000	F-NT2RP3004000	0.93	Yes	HRIFA025276a	0.93	Yes
NT2RP3004321	F-NT2RP3004321	0.81	Yes	HRIFA025786a	0.81	Yes
NT2RP3004355	F-NT2RP3004355	0.6	Yes	HRIFA025360a	0.6	Yes
NT2RP3004374	F-NT2RP3004374	0.58	Yes	HRIFA024533a	0.74	Yes
NT2RP4001001	F-NT2RP4001001	0.53	Yes	HRIFA009214a	0.5	Yes
NT2RP4002715	F-NT2RP4002715	0.94	Yes	HRIFA024921a	0.53	Yes
OVARC1000298	F-OVARC1000298	0.61	Yes	HRIFA004852a	0.59	Yes
OVARC1000775	F-OVARC1000775	0.7	Yes	HRIFA011347a	0.7	Yes
OVARC1000811	F-OVARC1000811	0.52	Yes	HRIFA000974a	0.39	Yes
OVARC1000853	F-OVARC1000853	0.94	Yes	HRIFA011403a	0.94	Yes
OVARC1001222	F-OVARC1001222	0.79	Yes	HRIFA022714a	0.67	Yes
OVARC1001807	F-OVARC1001807	0.52	Yes	HRIFA021069a	0.52	Yes
OVARC1001833	F-OVARC1001833	0.9	Yes	HRIFA021136a	0.9	Yes
PLACE1000231	F-PLACE1000231	0.52	Yes	HRIFA011802a	0.52	Yes
PLACE1000560	F-PLACE1000560	0.88	Yes	HRIFA012022a	0.88	Yes
PLACE1000740	F-PLACE1000740	0.57	Yes	HRIFA012151a	0.57	Yes

2/4 (continued)

Table 9

The 170 clones in which the selected clones have the maximal ATGpr1 score 0.5 or higher, and were predicted to contain the signal sequence by the PSORT

name of clone	name of sequence	maximal ATGpr1 score	signal	name of representative sequence	maximal ATGpr1 score	signal
PLACE1000912	F-PLACE1000912	0.9	Yes	HRIFA012282a	0.9	Yes
PLACE1000914	F-PLACE1000914	0.94	Yes	HRIFA012283a	0.94	Yes
PLACE1000927	F-PLACE1000927	0.71	Yes	HRIFA012290a	0.71	Yes
PLACE1000986	F-PLACE1000986	0.51	Yes	HRIFA012333a	0.51	Yes
PLACE1001100	F-PLACE1001100	0.76	Yes	HRIFA012417a	0.76	Yes
PLACE1001183	F-PLACE1001183	0.69	Yes	HRIFA012480a	0.69	Yes
PLACE1001229	F-PLACE1001229	0.65	Yes	HRIFA012513a	0.65	Yes
PLACE1001407	F-PLACE1001407	0.83	Yes	HRIFA012069a	0.94	Yes
PLACE1001788	F-PLACE1001788	0.6	Yes	HRIFA012881a	0.6	Yes
PLACE1002374	F-PLACE1002374	0.88	Yes	HRIFA013265a	0.92	Yes
PLACE1002518	F-PLACE1002518	0.94	Yes	HRIFA018849a	0.78	Yes
PLACE1003839	F-PLACE1003839	0.67	Yes	HRIFA014178a	0.6	Yes
PLACE1003845	F-PLACE1003845	0.92	Yes	HRIFA019185a	0.92	Yes
PLACE1004199	F-PLACE1004199	0.94	Yes	HRIFA014417a	0.94	Yes
PLACE1004282	F-PLACE1004282	0.94	Yes	HRIFA014467a	0.94	Yes
PLACE1004305	F-PLACE1004305	0.87	Yes	HRIFA014482a	0.87	Yes
PLACE1004637	F-PLACE1004637	0.89	Yes	HRIFA014692a	0.89	Yes
PLACE1005005	F-PLACE1005005	0.55	Yes	HRIFA014953a	0.55	Yes
PLACE1005250	F-PLACE1005250	0.52	Yes	HRIFA015129a	0.52	Yes
PLACE1005410	F-PLACE1005410	0.61	Yes	HRIFA015236a	0.61	Yes
PLACE1005725	F-PLACE1005725	0.92	Yes	HRIFA015443a	0.92	Yes
PLACE1005768	F-PLACE1005768	0.62	Yes	HRIFA015471a	0.62	Yes
PLACE1005927	F-PLACE1005927	0.66	Yes	HRIFA015568a	0.94	Yes
PLACE1006079	F-PLACE1006079	0.56	Yes	HRIFA015671a	0.56	Yes
PLACE1006093	F-PLACE1006093	0.59	Yes	HRIFA015682a	0.59	Yes
PLACE1006219	F-PLACE1006219	0.94	Yes	HRIFA015764a	0.93	Yes
PLACE1006809	F-PLACE1006809	0.66	Yes	HRIFA016129a	0.66	Yes
PLACE1007040	F-PLACE1007040	0.87	Yes	HRIFA013288a	0.87	Yes
PLACE1007096	F-PLACE1007096	0.59	Yes	HRIFA012167a	0.82	Yes
PLACE1007626	F-PLACE1007626	0.67	Yes	HRIFA016623a	0.67	Yes
PLACE1007971	F-PLACE1007971	0.74	Yes	HRIFA016838a	0.74	Yes
PLACE1008985	F-PLACE1008985	0.65	Yes	HRIFA017457a	0.48	Yes
PLACE1009067	F-PLACE1009067	0.59	Yes	HRIFA017509a	0.59	Yes
PLACE1009196	F-PLACE1009196	0.73	Yes	HRIFA017594a	0.73	Yes
PLACE1009527	F-PLACE1009527	0.58	Yes	HRIFA017791a	0.87	Yes
PLACE1009982	F-PLACE1009982	0.94	Yes	HRIFA018075a	0.94	Yes
PLACE1011236	F-PLACE1011236	0.52	Yes	HRIFA018827a	0.66	Yes
PLACE2000219	F-PLACE2000219	0.73	Yes	HRIFA034010a	0.73	Yes
SKNMC1000004	F-SKNMC1000004	0.94	Yes	HRIFA030097a	0.94	Yes
THYRO1000036	F-THYRO1000036	0.83	Yes	HRIFA027754a	0.83	Yes
THYRO1000099	F-THYRO1000099	0.94	Yes	HRIFA027803a	0.94	Yes
THYRO1001237	F-THYRO1001237	0.94	Yes	HRIFA030248a	0.94	Yes
THYRO1001327	F-THYRO1001327	0.93	Yes	HRIFA025125a	0.94	Yes
THYRO1001495	F-THYRO1001495	0.89	Yes	HRIFA030394a	0.89	Yes
THYRO1001523	F-THYRO1001523	0.71	Yes	HRIFA030408a	0.71	Yes
THYRO1001725	F-THYRO1001725	0.94	Yes	HRIFA029107a	0.94	Yes
Y79AA1000226	F-Y79AA1000226	0.94	Yes	HRIFA027874a	0.94	Yes
Y79AA1000521	F-Y79AA1000521	0.92	Yes	HRIFA027961a	0.92	Yes
Y79AA1000776	F-Y79AA1000776	0.78	Yes	HRIFA028401a	0.78	Yes
Y79AA1000959	F-Y79AA1000959	0.9	Yes	HRIFA028465a	0.9	Yes
Y79AA1001013	F-Y79AA1001013	0.94	Yes	HRIFA011193a	0.94	Yes
Y79AA1001264	F-Y79AA1001264	0.94	Yes	HRIFA028573a	0.94	Yes
Y79AA1001328	F-Y79AA1001328	0.91	Yes	HRIFA028592a	0.91	Yes
Y79AA1001427	F-Y79AA1001427	0.65	Yes	HRIFA028652a	0.65	Yes

3/4 (continued)

EP 1 130 094 A2

Table 10

The 170 clones in which the selected clones have the maximal ATGpr1 score 0.5 or higher, and were predicted to contain the signal sequence by the PSORT

name of clone	name of sequence	maximal ATGpr1 score	signal	name of representative sequence	maximal ATGpr1 score	signal
Y79AA1001430	F-Y79AA1001430	0.94	Yes	HRIFA028654a	0.94	Yes
Y79AA1001530	F-Y79AA1001530	0.94	Yes	HRIFA010206a	0.94	Yes
Y79AA1001592	F-Y79AA1001592	0.94	Yes	HRIFA028708a	0.94	Yes
Y79AA1001793	F-Y79AA1001793	0.89	Yes	HRIFA032066a	0.89	Yes
Y79AA1001795	F-Y79AA1001795	0.59	Yes	HRIFA032067a	0.59	Yes
Y79AA1001863	F-Y79AA1001863	0.56	Yes	HRIFA032097a	0.15	Yes
Y79AA1002022	F-Y79AA1002022	0.94	Yes	HRIFA033718a	0.94	Yes
Y79AA1002373	F-Y79AA1002373	0.79	Yes	HRIFA032271a	0.79	Yes

4/4

Table 11

The 35 clones in which the selected clones have the maximal ATGpr1 score 0.3 or higher and less than 0.5, and were predicted to contain the signal sequence by the PSORT

name of clone	name of sequence	maximal ATGpr1 score	signal	representative sequence	maximal ATGpr1 score	signal
HEMBA1000907	F-HEMBA1000907	0.39	Yes	HRIFA000845a	0.94	Yes
HEMBA1003602	F-HEMBA1003602	0.43	Yes	HRIFA020109a	0.43	Yes
HEMBA1004797	F-HEMBA1004797	0.45	Yes	HRIFA020883a	0.66	Yes
HEMBB1000447	F-HEMBB1000447	0.31	Yes	HRIFA001558a	0.76	Yes
MAMMA1000591	F-MAMMA1000591	0.34	Yes	HRIFA026303a	0.77	Yes
MAMMA1000681	F-MAMMA1000681	0.35	Yes	HRIFA026364a	0.94	Yes
MAMMA1000986	F-MAMMA1000986	0.37	Yes	HRIFA021611a	0.37	Yes
MAMMA1001893	F-MAMMA1001893	0.44	Yes	HRIFA027485a	0.9	Yes
MAMMA1001957	F-MAMMA1001957	0.48	Yes	HRIFA027536a	0.94	Yes
NT2RM2001941	F-NT2RM2001941	0.44	Yes	HRIFA032011a	0.94	Yes
NT2RP1000050	F-NT2RP1000050	0.47	Yes	HRIFA005102a	0.54	Yes
NT2RP1000903	F-NT2RP1000903	0.38	Yes	HRIFA005650a	0.38	Yes
NT2RP2003469	F-NT2RP2003469	0.33	Yes	HRIFA008661a	0.9	Yes
NT2RP2003664	F-NT2RP2003664	0.36	Yes	HRIFA008790a	0.89	Yes
NT2RP2004447	F-NT2RP2004447	0.36	Yes	HRIFA009339a	0.93	Yes
NT2RP2006042	F-NT2RP2006042	0.37	Yes	HRIFA010425a	0.69	Yes
NT2RP3001195	F-NT2RP3001195	0.44	Yes	HRIFA023227a	0.94	Yes
NT2RP3003354	F-NT2RP3003354	0.3	Yes	HRIFA008212a	0.51	Yes
NT2RP3003469	F-NT2RP3003469	0.3	Yes	HRIFA025143a	0.3	Yes
NT2RP3003963	F-NT2RP3003963	0.44	Yes	HRIFA008949a	0.62	Yes
NT2RP3004133	F-NT2RP3004133	0.35	Yes	HRIFA025706a	0.94	Yes
NT2RP3004309	F-NT2RP3004309	0.4	Yes	HRIFA025778a	0.92	Yes
OVARC1000208	F-OVARC1000208	0.4	Yes	HRIFA010942a	0.4	Yes
PLACE1001536	F-PLACE1001536	0.33	Yes	HRIFA012761a	0.31	Yes
PLACE1003407	F-PLACE1003407	0.48	Yes	HRIFA013899a	0.48	Yes
PLACE1003428	F-PLACE1003428	0.37	Yes	HRIFA013911a	0.61	Yes
PLACE1003460	F-PLACE1003460	0.38	Yes	HRIFA013932a	0.94	Yes
PLACE1005569	F-PLACE1005569	0.32	Yes	HRIFA015351a	0.68	Yes
PLACE1006277	F-PLACE1006277	0.32	Yes	HRIFA015802a	0.65	Yes
PLACE1010251	F-PLACE1010251	0.32	Yes	HRIFA018238a	0.62	Yes
THYRO1000196	F-THYRO1000196	0.44	Yes	HRIFA029050a	0.71	Yes
THYRO1000795	F-THYRO1000795	0.33	Yes	HRIFA029327a	0.92	Yes
THYRO1000999	F-THYRO1000999	0.4	Yes	HRIFA030203a	0.4	Yes
THYRO1001478	F-THYRO1001478	0.47	Yes	HRIFA030385a	0.89	Yes
Y79AA1000426	F-Y79AA1000426	0.47	Yes	HRIFA027940a	0.92	Yes

Table 12

41 clones from which selected clones have the maximal ATGpr1 score 0 or higher and less than 0.3, and predicted to be containing the signal sequence by the PSORT

name of clone	name of sequence	maximal ATGpr1 score	signal	representative sequence	maximal ATGpr1 score	signal
HEMBA1000300	F-HEMBA1000300	0.13	Yes	HRIFA000284a	0.13	Yes
HEMBA1002164	F-HEMBA1002164	0.11	Yes	HRIFA001972a	0.74	Yes
HEMBA1002239	F-HEMBA1002239	0.17	Yes	HRIFA002037a	0.17	Yes
HEMBA1002421	F-HEMBA1002421	0.22	Yes	HRIFA005392a	0.9	Yes
HEMBA1003294	F-HEMBA1003294	0.15	Yes	HRIFA020163a	0.15	Yes
HEMBA1006572	F-HEMBA1006572	0.06	Yes	HRIFA021543a	0.62	Yes
HEMBA1007013	F-HEMBA1007013	0.19	Yes	HRIFA021906a	0.82	Yes
HEMBA1000567	F-HEMBA1000567	0.09	Yes	HRIFA029730a	0.15	Yes
HEMBA1002663	F-HEMBA1002663	0.29	Yes	HRIFA031871a	0.29	Yes
MAMMA1001043	F-MAMMA1001043	0.17	Yes	HRIFA026764a	0.81	Yes
MAMMA1001284	F-MAMMA1001284	0.26	Yes	HRIFA026889a	0.26	Yes
MAMMA1001901	F-MAMMA1001901	0.17	Yes	HRIFA027493a	0.17	Yes
MAMMA1002224	F-MAMMA1002224	0.13	Yes	HRIFA027717a	0.13	Yes
NT2RM2000306	F-NT2RM2000306	0.25	Yes	HRIFA021985a	0.25	Yes
NT2RM2000410	F-NT2RM2000410	0.22	Yes	HRIFA022055a	0.82	Yes
NT2RP2000479	F-NT2RP2000479	0.24	Yes	HRIFA000822a	0.12	Yes
NT2RP2001495	F-NT2RP2001495	0.19	Yes	HRIFA007228a	0.78	Yes
NT2RP2001948	F-NT2RP2001948	0.29	Yes	HRIFA007565a	0.89	Yes
NT2RP3000645	F-NT2RP3000645	0.2	Yes	HRIFA022890a	0.91	Yes
NT2RP3003076	F-NT2RP3003076	0.23	Yes	HRIFA024978a	0.65	Yes
NT2RP4001879	F-NT2RP4001879	0.26	Yes	HRIFA017818a	0.79	Yes
NT2RP4002451	F-NT2RP4002451	0.11	Yes	HRIFA018447a	0.34	Yes
OVARC1000439	F-OVARC1000439	0.15	Yes	HRIFA011105a	0.21	Yes
OVARC1001727	F-OVARC1001727	0.22	Yes	HRIFA019960a	0.22	Yes
PLACE1002080	F-PLACE1002080	0.17	Yes	HRIFA013092a	0.76	Yes
PLACE1002095	F-PLACE1002095	0.23	Yes	HRIFA013103a	0.61	Yes
PLACE1004028	F-PLACE1004028	0.12	Yes	HRIFA014303a	0.12	Yes
PLACE1004482	F-PLACE1004482	0.13	Yes	HRIFA014590a	0.57	Yes
PLACE1005383	F-PLACE1005383	0.11	Yes	HRIFA015219a	0.52	Yes
PLACE1005544	F-PLACE1005544	0.08	Yes	HRIFA009852a	0.41	Yes
PLACE1005660	F-PLACE1005660	0.2	Yes	HRIFA015409a	0.2	Yes
PLACE1006443	F-PLACE1006443	0.27	Yes	HRIFA015902a	0.89	Yes
PLACE1007296	F-PLACE1007296	0.22	Yes	HRIFA016430a	0.27	Yes
PLACE1008469	F-PLACE1008469	0.27	Yes	HRIFA017146a	0.94	Yes
PLACE1008984	F-PLACE1008984	0.11	Yes	HRIFA017456a	0.74	Yes
PLACE4000455	F-PLACE4000455	0.23	Yes	HRIFA012333a	0.51	Yes
SKNMC1000014	F-SKNMC1000014	0.15	Yes	HRIFA030106a	0.76	Yes
THYRO1001702	F-THYRO1001702	0.14	Yes	HRIFA030511a	0.8	Yes
Y79AA1000270	F-Y79AA1000270	0.21	Yes	HRIFA005644a	0.63	Yes
Y79AA1001056	F-Y79AA1001056	0.27	Yes	HRIFA028497a	0.27	Yes
Y79AA1001803	F-Y79AA1001803	0.08	Yes	HRIFA032073a	0.68	Yes

Table 13

Four clones from which selected clones have the maximal ATGpr1 score 0.5 or higher, and predicted to be lacking the signal sequence by the PSORT

name of clone	name of sequence	maximal ATGpr1 score	signal	name of representative sequence	maximal ATGpr1 score	signal
NT2RP3002281	F-NT2RP3002281	0.81	No	HRIFA012999a	0.61	Yes
NT2RP3002721	F-NT2RP3002721	0.94	No	HRIFA023305a	0.57	Yes
NT2RP3004083	F-NT2RP3004083	0.94	No	HRIFA008387a	0.76	Yes
PLACE1005669	F-PLACE1005669	0.94	No	HRIFA012513a	0.65	Yes

Table 14

Two clones from which selected clones have the maximal ATGpr1 score 0.3 or higher and less than 0.5 and predicted to have no signal sequence by the PSORT

name of clone	name of sequence	maximal ATGpr1 score	signal	representative sequence	maximal ATGpr1 score	signal
NT2RP3000481	F-NT2RP3000481	0.47	No	HRIFA028614a	0.93	Yes
NT2RP3003559	F-NT2RP3003559	0.48	No	HRIFA025514a	0.45	Yes

Table 15

Two clones from which selected clones have the maximal ATGpr1 score 0 or higher and less than 0.3, and predicted to have no signal sequence by the PSORT

name of clone	name of sequence	maximal ATGpr1 score	signal	representative sequence	maximal ATGpr1 score	signal
PLACE1005601	F-PLACE1005601	0.12	No	HRIFA010593a	0.64	Yes
PLACE1006786	F-PLACE1006786	0.22	No	HRIFA012333a	0.51	Yes

EXAMPLE 12

A method for the selection of clones based on the top hit data in the homology search against the SwissProt (a method for the selection of a novel full-length protein that is predicted to have a function based on the top hit data).

[0176] The representative sequences of the clusters were discarded in which the 5'-end sequence is identical (90% or more matching), or not similar (the compared part contains a sequence of 25 nucleotides or shorter and the similarity is lower than 60%) to the top hit data in the SwissProt. Then, the remaining representative sequences which has similarity to the representative sequences of the 5'-ends were classified by a group of the above keywords (some representative sequences belong to a group by multiple keywords), and then clones were selected from the clusters. The names and the corresponding SEQ IDs of the representative sequences, and also the names of the introns (including the representative sequence of the 5'-end or the 3'-end, or ESTs) used for selecting the clones from the representative sequences and the corresponding SEQ IDs are shown in the last part of the present specification (Table 368 and 369, respectively). HRIFA indicates the representative sequence of the 5'-end group, and HRIRA indicates the representative sequence of the 3'-end group.

[0177] In principle, from the clusters containing only a single clone (the 5'-end sequence is the representative sequence of the cluster), the clone was selected. However, in the following cases, the clone containing a longer 5'-end was selected: where the maximal ATGpr1 score was less than 0.5, the 5'-end sequence of the clone to be selected was not complete, or the 5'-end of the clone was found to be obviously short nevertheless the clone should not be

included in the same cluster based on the BLAST analysis between the other clones, or further, the 5'-end sequence of the said clone, which corresponds to the 3'-ends of the other clones belonging to the same cluster in which the 3'-end of the said clone was included, was turn out to be longer than those of the other clones by assembling them. When there were two clones in the same cluster, judged to be full-length, and it was difficult to determine which clone has the longer 5'-end even by assembling them, all the clones were selected. As a result, the ATGpr1 score in some clones became less than 0.5 or less than 0.3. The fullness ratio of these clones became lower, but there is still a possibility that the clones are full-length.

[0178] In the case in which multiple clones were contained in a cluster, selected was the clone having the longest 5'-end in the clones judged to be full-length compared to the human mRNA or human EST. However, in the following cases, the clone containing a longer 5'-end was selected: where the maximal ATGpr1 score was less than 0.5, the 5'-end sequence of the clone to be selected was not complete, or the 5'-end of the clone was found to be obviously short nevertheless the clone should not be included in the same cluster based on the BLAST analysis between the other clones, or further, the 5'-end sequence of the said clone, which corresponds to the 3'-ends of the other clones belonging to the same cluster in which the 3'-end of the said clone was included, was turn out to be longer than those of the other clones by assembling them. When there were two clones in the same cluster, judged to be full-length, and it was difficult to determine which clone has the longer 5'-end even by assembling them, all the clones were selected. As a result, the ATGpr1 score in some clones became less than 0.5 or less than 0.3. These clones can still be full-length.

[0179] Based on the top hit data in the SwissProt homology search, 658 clones were selected. Among them, 446 clones were selected by the keywords, secretion or membrane. Using the keyword, glycoprotein, 243 clones were selected. 51 clones were selected by the keywords for signal transduction. With the keywords for transcription, 130 clones were selected. 17 clones were selected by the keywords for disease.

[0180] Among the 446 clones selected by the keywords, secretion or membrane, 77 clones were overlapped with those selected by combining the ATGpr1 score and prediction by the PSORT for the existence of a signal sequence. Also, many clones were overlapped with those selected by the keyword, glycoprotein. Moreover, some clones were overlapped with the clones selected by the keywords for diseases.

[0181] Among the 243 clones selected by the keyword, glycoprotein, 53 clones were overlapped with those selected by combining the ATGpr1 score and prediction by the PSORT for the existence of a signal sequence. Also, many clones were overlapped with those selected by the keywords, secretion or membrane. Moreover, some clones were overlapped with the clones selected by the keywords in diseases.

[0182] Among the clones selected by the top hit data in the homology search on the SwissProt, 532 clones were having the maximal ATGpr1 score 0.5 or higher. 59 clones were having the maximal score 0.3 or higher and less than 0.5. 67 clones were with the maximal score less than 0.3.

[0183] When the maximal ATGpr1 score is 0.5 or higher, 0.3 or higher, no less than 0, the expected fullness ratio is as shown in Table 3, 4, 5, and 6, respectively.

Table 16

The representative sequences of the most homologous sequences in the SwissProt with the keyword(s) "growth factor", "cytokine", or "hormone", and the selected clones.	
name of clone	name of representative sequence
HEMBA1001563	HRIFA001439a
HEMBA1003047	HRIFA002743a
HEMBA1005070	HRIFA020144a
HEMBA1006724	HRIFA021620a
HEMBA1006916	HRIFA021855a
MAMMA1001066	HRIFA027355a
MAMMA1001634	HRIFA027187a
MAMMA1002165	HRIFA027673a
NT2RM4000326	HRIFA032530a
NT2RM4001377	HRIFA005300a
NT2RP2000447	HRIFA006448a
NT2RP2000663	HRIFA006609a
NT2RP2000903	HRIFA006798a
NT2RP2002974	HRIFA027860a
NT2RP2003369	HRIFA008596a
NT2RP2004141	HRIFA009123a

EXAMPLE 13

Selection of cDNA clone NT2RP2036580

- 5 **[0184]** Clone NT2RP2006580 as well as clone HEMBA1000121 was selected from the representative sequences belonging to HRIFA000116a cluster of the most homologous sequence in the SwissProt with the keywords "trans-membrane". Although each of the clones, HEMBA1000121 and NT2RP2006580, was assembled with other clones for 5' extension, any other clones did not extend the clones toward the 5' direction. Accordingly, it is possible that both clones are full-length cDNA clones. The maximal ATGpr1 score of F-NT2RP2006580 is 0.37, and therefore, the fullness ratio is low. However, it is still possible for the sequence to cover the full-length.
- 10 **[0185]** Thus, the total number of selected clones is 830. Based on the top matching data resulted from Swiss-Prot homology search, 659 clones were selected. From them, 447 clones were selected by the keywords of "secretion" and "membrane". Among the clones selected based on the top matching data, 60 clones exhibited the maximal ATGpr1 score of 0.3 or higher and less than 0.5.
- 15 **[0186]** The sequences of F-NT2RP2006580 and R-NT2RP2006580 are shown in SEQ ID NO: 2545 and SEQ ID NO: 2546, respectively.

EXAMPLE 14

20 Full-length sequence analysis and homology search

- [0187]** Full-length sequence was determined for each selected cDNA clones. The nucleotide sequence determination was performed mainly by the dye-terminator method using custom synthesized DNA primers according to the primer walking procedure (custom synthesized DNA primers were used for sequencing; sequencing reaction was performed with DNA sequencing reagent supplied by PE Biosystems according to the supplier's manual; and the samples were analyzed in an automatic sequencer made by the same supplier). Sequence determination of some clones was carried out in the same manner but using a Licor DNA sequencer. Overlapping partial nucleotide sequences, which were obtained by the above-described method, were assembled together to determine a full-length nucleotide sequence. Amino acid sequences were then deduced from the determined full-length nucleotide sequences. However, amino acid sequence is not shown for a clone of which coding region was hard to be deduced or of which amino acid sequence has less than 100 amino acid residues. SEQ ID NOs corresponding to the respective clones are indicated in Table 370.
- 25 **[0188]** GenBank, Swiss-Prot and UniGene were searched for the determined nucleotide sequences by BLAST analysis. Matching data of cDNA clone which exhibits higher homology and of which functions are easily predicted based on the nucleotide sequences and the deduced amino acid sequences are selected from the BLAST analysis matching data with P value of 10^{-4} or less. The matching data selected are listed herein. However, there are some clones that did not match the criteria for judgment and such matching data of BLAST analysis are not shown herein. The results of homology search indicated in the last part of this specification are as follows.
- 30 **[0189]** Homology search result 1: data obtained by the homology search of Swiss-Prot database for representative sequences of the 5'-end cluster
- 40 **[0190]** Homology search result 2: homology of representative sequences of the 5'-end cluster to the data in Swiss-Prot database; the P value is 10^{-10} or less
- [0191]** Homology search result 3: homology of representative sequences of the 5'-end cluster to the data in Swiss-Prot database; the P value is higher than 10^{-10} and 10^{-4} or less
- 45 **[0192]** Homology search result 4: homology of representative sequences of the 5'-end cluster to the data in Swiss-Prot database; the P value is higher than 10^{-4} and 1 or less
- [0193]** Homology search result 5: data obtained by the homology search of Swiss-Prot database for 5'-end sequences of cDNA clone
- [0194]** Homology search result 6: data obtained by the homology search of GenBank database (<http://www.ncbi.nlm.nih.gov/web/GenBank/>) except for EST and STS sequence data for 5'-end sequences of cDNA clone
- 50 **[0195]** Homology search result 7: data obtained by the homology search of GenBank database (<http://www.ncbi.nlm.nih.gov/web/GenBank/>) except for EST and STS sequence data for 3'-end sequences of cDNA clone
- [0196]** Homology search result 8: data obtained by the homology search of Human UniGene database (<http://www.ncbi.nlm.nih.gov/Unigene/>) for 5'-end sequences of cDNA clone
- 55 **[0197]** Homology search result 9: data obtained by the homology search of Human UniGene database (<http://www.ncbi.nlm.nih.gov/Unigene/>) for 3'-end sequences of cDNA clone
- [0198]** Homology search result 10: result obtained by the homology search for full-length nucleotide sequences and deduced amino acid sequences
- [0199]** The P value indicates similarity between two sequences as a score by considering the probability that the two

sequences are accidentally similar. In general, as the value is lower, the similarity is higher. In general, as the value is lower, the homology is higher (Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.L. (1990) "Basic local alignment search tool." J. Mol. Biol. 215:403-410; Gish, W. & States, D.J. (1993) "Identification of protein coding regions by database similarity search." Nature Genet. 3:266-272).

5

Example 15. Gene expression analysis with hybridization using high density DNA filter

[0200] Nylon membrane for DNA spotting was prepared according to the following procedure. E. coli was cultured in each well of a 96-well plate (in a LB medium at 37. for 16 hours). A sample of each culture was suspended in 10 . 1 of sterile water in a well of a 96-well plate. The plate was heated at 100. for 10 minutes. Then, the boiled samples were analyzed by PCR. PCR was performed in a 20 .1 solution by using TaKaRa PCR Amplification Kit (Takara) according to the supplier's protocol. Primers used for the amplification of an insert cDNA in a plasmid were a pair of sequencing primers, ME761FW (5' tacggaagtgttactctgc 3' / SEQ ID NO: 3591) and ME1250RV (5' tgtgggaggttttctcta 3' / SEQ ID NO: 3592), or a pair of primers, M13M4 (5' gtttccagtcacgac 3' / SEQ ID NO: 3593) and M13RV (5' cag-gaaacagctatgac 3' / SEQ ID NO: 3594). PCR was performed using a thermal cycler, GeneAmp System 9600 (PE Biosystems) at 95. for 5 minutes; at 95. for 10 seconds and at 68. for 1 minute for 10 cycles; at 98. for 20 seconds and at 60. for 3 minutes for 20 cycles; and at 72. for 10 minutes. After the PCR, the 20 .1 reaction solution was loaded onto a 1% agarose gel and fractionated by electrophoresis. DNA on the gel was stained with ethidium bromide to confirm the amplification of cDNA. When cDNAs were not amplified by PCR, plasmids containing the corresponding insert cDNAs were prepared by the alkali-extraction method (J. Sambrook, E.F., Fritsh, & T. Maniatis, "Molecular Cloning, A laboratory manual/ 2nd edition, Cold Spring Harbor Laboratory Press, 1989).

[0201] Preparation of DNA array was carried out by the following procedure. A sample of a DNA solution was added in each well of a 384-well plate. DNA was spotted onto a nylon membrane (Boehringer) by using a 384-pin tool of Biomek 2000 Laboratory Automation System (Beckman-Coulter). Specifically, the 384-well plate containing the DNA was placed under the 384-pin tool. The independent 384 needles were simultaneously dipped into the DNA solution for DNA deposition. The needles were gently pressed onto a nylon membrane and the DNA deposited at the tips of needles was spotted onto the membrane. Denaturation of the spotted DNA and immobilization of the DNA on the nylon membrane were carried out according to standard methods (J. Sambrook, E.F., Fritsh, & T. Maniatis, "Molecular Cloning, A laboratory manual/ 2nd edition, Cold Spring Harbor Laboratory Press, 1989).

[0202] A probe for hybridization was radioisotope-labeled first strand cDNA. Synthesis of the first strand cDNA was performed by using Thermoscript^(TM) RT-PCR System (GIBCO). Specifically, the first strand cDNA was synthesized by using 1.5 .g of mRNAs from various human tissues (Clontech), 1 .1 of 50.M Oligo(dT)20 and 50.Ci [³³P]dATP according to an attached protocol. Purification of a probe was carried out by using ProbeQuant^(TM) G-50 micro column (Amersham-Pharmacia Biotech) according to an attached protocol. In the next step, 2 units of E. coli RNase H were added to the reaction mixture. The mixture was incubated at room temperature for 10 minutes, and then, 100.g of human COT-1 DNA (GIBCO) was added thereto. The mixture was incubated at 97. for 10 minutes and then was allowed to stand on ice to give hybridization probe.

[0203] Hybridization of the radioisotope-labeled probe to the DNA array was performed according to standard methods (J. Sambrook, E.F., Fritsh, & T. Maniatis, Molecular Cloning, A laboratory manual/ 2nd edition, Cold Spring Harbor Laboratory Press, 1989). The membrane was washed as follows: the nylon membrane was washed 3 times by incubating it in Washing solution 1 (2xSSC, 1% SDS) at room temperature (about 26.) for 20 minutes; then the membrane was washed 3 times by incubating it in Washing solution 2 (0.1xSSC, 1% SDS) at 65. for 20 minutes.

[0204] Autoradiography was performed by using an image plate for BAS2000 (Fuji Photo Film Co., Ltd.). Specifically, the nylon membrane with probe hybridized thereon was wrapped with a piece of Saran Wrap and brought into tight contact with the image plate on the light-sensitive surface. The membrane with the image plate was placed in an imaging cassette for radioisotope and allowed to stand in dark place for 4 hours. The radioactivity recorded on the image plate was analyzed by using BAS2000 (Fuji Photo Film Co., Ltd.). The activity was subjected to electronic conversion and recorded as an image file of autoradiogram. The signal intensity of each DNA spot was analyzed by using Visage High Density Grid Analysis Systems (Genomic Solutions Inc.). The signal intensity was converted into numerical data. The data were taken in duplicate. The reproducibility was assessed by comparing the signal intensities of the corresponding spots on the duplicated DNA filters that were hybridized to a single DNA probe (Figure 2). In 95% of entire spots, the ratio between the corresponding spots falls within a range of 2 or less, and the correlation coefficient is $r=1.97$. Thus, the reproducibility is satisfactory.

[0205] The detection sensitivity in gene expression analysis was estimated by examining increases in the signal intensity of probe concentration-dependent spot in hybridization using a probe complementary to the DNA spotted on the nylon membrane. DNA used was PLACE 1008092 (the same as DNA deposited in GenBank under an Accession No. AF107253). The DNA array with DNA of PLACE1008092 was prepared according to the above-mentioned method. The probe used was prepared as follows: mRNA was synthesized in vitro from the clone, PLACE1008092. By using

5 this mRNA as a template, radioisotope-labeled first strand cDNA was synthesized in the same manner as described above, and the cDNA was used as the probe. In order to synthesize mRNA from PLACE1008092 in vitro, a plasmid in which the 5' end of the cDNA PLACE1008092 was ligated to the T7 promoter of pBluescript SK(-) was constructed. Specifically, the PLACE1008092 insert was cut out from pME18SFL3 carrying the cDNA at a DralII site thereof by XhoI digestion. The resulting PLACE1008092 fragment was ligated to XhoI-predigested pBluescript SK(-) by using DNA ligation kit ver.2 (Takara). The in vitro mRNA synthesis from PLACE1008092 inserted into pBluescript SK(-) was carried out by using Ampliscribe^(TM) T7 high yield transcription kit (Epicentre technologies). Hybridization and the analysis of signal intensity of each DNA spot were performed by the same methods as described above. When the probe concentration is 1×10^7 .g/ml or less, there was no increase of signal intensity proportional to the probe concentration. Therefore, it was assumed to be difficult to compare the signals with one another in this concentration range. Thus, the spots with the intensity of 40 or less were uniformly taken as low level signals (Figure 3). Within a concentration of the probe ranging from 1×10^7 .g/ml to 0.1.g/ml, the signal was found to increase in a probe concentration-dependent manner. The detection limit represented as the ratio of the expression level of test mRNA to that of total mRNA in a sample was 1:100,000.

15 **[0206]** Tables 28-184 (also containing clones without description in Examples) show the expression of each cDNA in human normal tissues (heart, lung, pituitary gland, thymus, brain, kidney, liver and spleen). The expression levels are indicated with numerical values of 0-10,000. Genes that were expressed in at least a single tissue are indicated below by the corresponding clone names:

	clone:	BNGH41000020,	BNGH41000087,	BNGH41000091,	HEMBA1000121,	HEMBA1000275,
20	HEMBA1000300,	HEMBA1000443,	HEMBA1000462,	HEMBA1000477,	HEMBA1000634,	HEMBA1000713,
	HEMBA1000835,	HEMBA1000875,	HEMBA1000940,	HEMBA1000962,	HEMBA1001228,	HEMBA1001296,
	HEMBA1001390,	HEMBA1001563,	HEMBA1001621,	HEMBA1002048,	HEMBA1002131,	HEMBA1002163,
	HEMBA1002164,	HEMBA1002167,	HEMBA1002178,	HEMBA1002195,	HEMBA1002227,	HEMBA1002239,
	HEMBA1002316,	HEMBA1002421,	HEMBA1002524,	HEMBA1002551,	HEMBA1002767,	HEMBA1002985,
25	HEMBA1002992,	HEMBA1003047,	HEMBA1003072,	HEMBA1003101,	HEMBA1003120,	HEMBA1003230,
	HEMBA1003294,	HEMBA1003315,	HEMBA1003392,	HEMBA1003399,	HEMBA1003487,	HEMBA1003530,
	HEMBA1003945,	HEMBA1004007,	HEMBA1004067,	HEMBA1001085,	HEMBA1004110,	HEMBA1004391,
	HEMBA1004444,	HEMBA1004454,	HEMBA1004505,	HEMBA1004797,	HEMBA1004952,	HEMBA1005070,
	HEMBA1005084,	HEMBA1005145,	HEMBA1005230,	HEMBA1005246,	HEMBA1005337,	HEMBA1005430,
30	HEMBA1005449,	HEMBA1005489,	HEMBA1005545,	HEMBA1005698,	HEMBA1005929,	HEMBA1005945,
	HEMBA1005016,	HEMBA1006171,	HEMBA1006276,	HEMBA1006311,	HEMBA1006335,	HEMBA1006357,
	HEMBA1006430,	HEMBA1006482,	HEMBA1006517,	HEMBA1006544,	HEMBA1006658,	HEMBA1006707,
	HEMBA1006749,	HEMBA1006770,	HEMBA1006902,	HEMBA1006912,	HEMBA1006916,	HEMBA1006960,
	HEMBA1007013,	HEMBA1007057,	HEMBA1007063,	HEMBA1007291,	HEMBA1007332,	HEMBA1007106,
35	HEMBA1000309,	HEMBA1000447,	HEMBA1000542,			
	HEMBA1000567,	HEMBA1000642,	HEMBA1000905,	HEMBA1001026,	HEMBA1001048,	HEMBA1001407,
	HEMBA1001530,	HEMBA1001573,	HEMBA1001847,	HEMBA1001959,	HEMBA1001978,	HEMBA1002039,
	HEMBA1002041,	HEMBA1002051,	HEMBA1002162,	HEMBA1002228,	HEMBA1002302,	HEMBA1002427,
	HEMBA1002465,	HEMBA1002661,	HEMBA1002663,	HEMBA1002693,	MAMMA1000046,	MAMMA1000102,
40	MAMMA1000106,	MAMMA1000118,	MAMMA1000204,	MAMMA1000226,	MAMMA1000403,	MAMMA1000449,
	MAMMA1000457,	MAMMA1000473,	MAMMA1000528,	MAMMA1000591,	MAMMA1000614,	MAMMA1000652,
	MAMMA1000681,	MAMMA1000706,	MAMMA1000788,	MAMMA1000810,	MAMMA1000814,	MAMMA1000881,
	MAMMA1000986,	MAMMA1000994,	MAMMA1001043,	MAMMA1001066,	MAMMA1001094,	MAMMA1001141,
	MAMMA1001150,	MAMMA1001284,	MAMMA1001310,	MAMMA1001344,	MAMMA1001418,	MAMMA1001532,
45	MAMMA1001609,	MAMMA1001615,	MAMMA1001634,	MAMMA1001893,	MAMMA1001901,	MAMMA1001957,
	MAMMA1002070,	MAMMA1002091,	MAMMA1002095,	MAMMA1002128,	MAMMA1002142,	MAMMA1002165,
	MAMMA1002205,	MAMMA1002224,	MAMMA1002586,	MAMMA1003126,	NT2RM1000407,	NT2RM1000462,
	NT2RM1000542,	NT2RM1000789,	NT2RM1000855,	NT2RM1000858,	NT2RM2000241,	NT2RM2000306,
	NT2RM2000410,	NT2RM2000423,	NT2RM2000497,	NT2RM2000514,	NT2RM2000565,	NT2RM2000582,
50	NT2RM2000589,	NT2RM2000622,	NT2RM2000773,	NT2RM2001126,	NT2RM2001626,	NT2RM2001792,
	NT2RM2001941,	NT2RM4000198,	NT2RM4000295,	NT2RM4000444,	NT2RM4000593,	NT2RM4000761,
	NT2RM4000965,	NT2RM4000997,	NT2RM4001321,	NT2RM4001325,		
	NT2RM4001377,	NT2RM4001735,	NT2RM4001768,	NT2RM4001843,	NT2RP1000002,	NT2RP1000181,
	NT2RP1000271,	NT2RP1000300,	NT2RP1000325,	NT2RP1000465,	NT2RP1000468,	NT2RP1000740,
55	NT2RP1000903,	NT2RP1000981,	NT2RP2000092,	NT2RP2000178,	NT2RP2000240,	NT2RP2000447,
	NT2RP2000479,	NT2RP2000533,	NT2RP2000610,	NT2RP2000616,	NT2RP2000694,	NT2RP2000739,
	NT2RP2001200,	NT2RP2001223,	NT2RP2001388,	NT2RP2001469,	NT2RP2001480,	NT2RP2001514,
	NT2RP2001529,	NT2RP2001538,	NT2RP2001562,	NT2RP2001662,	NT2RP2001878,	NT2RP2001903,

EP 1 130 094 A2

	NT2RP2001921,	NT2RP2001956,	NT2RP2002015,	NT2RP2002063,	NT2RP2002188,	NT2RP2002232,
	NT2RP2002409,	NT2RP2002510,	NT2RP2002527,	NT2RP2002533,	NT2RP2002564,	NT2RP2002721,
	NT2RP2002824,	NT2RP2002942,	NT2RP2002974,	NT2RP2003138,	NT2RP2003179,	NT2RP2003210,
	NT2RP2003302,	NT2RP2003369,	NT2RP2003383,	NT2RP2003390,	NT2RP2003469,	NT2RP2003593,
5	NT2RP2003599,	NT2RP2003655,	NT2RP2003940,	NT2RP2003950,	NT2RP2004069,	NT2RP2004108,
	NT2RP2004141,	NT2RP2004179,	NT2RP2004205,	NT2RP2004447,	NT2RP2004524,	NT2RP2004556,
	NT2RP2004606,	NT2RP2004648,	NT2RP2004794,	NT2RP2004837,	NT2RP2004847,	NT2RP2005027,
	NT2RP2005069,	NT2RP2005163,	NT2RP2005181,	NT2RP2005247,	NT2RP2005378,	NT2RP2005391,
	NT2RP2005425,	NT2RP2005463,	NT2RP2005535,	NT2RP2005541,	NT2RP2005597,	NT2RP2005632,
10	NT2RP2005666,	NT2RP2005774,	NT2RP2005878,	NT2RP2005887,	NT2RP2005941,	NT2RP2006004,
	NT2RP2006042,	NT2RP2006092,	NT2RP2006099,	NT2RP2006269,	NT2RP3000011,	NT2RP3000022,
	NT2RP3000059,	NT2RP3000063,	NT2RP3000125,	NT2RP3000148,	NT2RP3000171,	NT2RP3000172,
	NT2RP3000201,	NT2RP3000232,	NT2RP3000304,	NT2RP3000378,	NT2RP3000436,	NT2RP3000460,
	NT2RP3000645,	NT2RP3000652,	NT2RP3000676,	NT2RP3000677,	NT2RP3000721,	NT2RP3000789,
15	NT2RP3000818,	NT2RP3000820,	NT2RP3000838,	NT2RP3000907,	NT2RP3000921,	NT2RP3001044,
	NT2RP3001159,	NT2RP3001170,	NT2RP3001195,	NT2RP3001271,	NT2RP3001388,	NT2RP3001560,
	NT2RP3001592,	NT2RP3001685,	NT2RP3001738,	NT2RP3001754,	NT2RP3001858,	NT2RP3001976,
	NT2RP3002015,	NT2RP3002160,	NT2RP3002281,	NT2RP3002311,	NT2RP3002324,	NT2RP3002353,
	NT2RP3002409,	NT2RP3002411,	NT2RP3002721,	NT2RP3002737,	NT2RP3002738,	NT2RP3002836,
20	NT2RP3002900,	NT2RP3002958,	NT2RP3003000,	NT2RP3003076,	NT2RP3003354,	NT2RP3003448,
	NT2RP3003469,	NT2RP3003473,	NT2RP3003532,	NT2RP3003614,	NT2RP3003729,	NT2RP3003849,
	NT2RP3003874,	NT2RP3003939,	NT2RP3003963,	NT2RP3004025,	NT2RP3004067,	NT2RP3004083,
	NT2RP3004090,	NT2RP3004119,	NT2RP3004130,	NT2RP3004133,	NT2RP3004202,	NT2RP3004294,
	NT2RP3004309,	NT2RP3004321,	NT2RP3004355,	NT2RP3004374,	NT2RP3004406,	NT2RP3004481,
25	NT2RP3004552,	NT2RP3004557,	NT2RP3004625,	NT2RP3004640,	NT2RP3004647,	NT2RP4000108,
	NT2RP4000634,	NT2RP4001877,	NT2RP4001879,	NT2RP4002187,	NT2RP4002715,	NT2RP4002750,
	OVARC1000090,	OVARC1000105,	OVARC1000137,	OVARC1000208,	OVARC1000255,	OVARC1000313,
	OVARC1000331,	OVARC1000410,	OVARC1000439,	OVARC1000467,	OVARC1000529,	OVARC1000553,
	OVARC1000775,	OVARC1000853,	OVARC1000873,	OVARC1000916,	OVARC1000956,	OVARC1000995,
30	OVARC1001030,	OVARC1001049,	OVARC1001086,	OVARC1001163,	OVARC1001260,	OVARC1001336,
	OVARC1001569,	OVARC1001570,	OVARC1001596,	OVARC1001807,	OVARC1001833,	OVARC1001991,
	PLACE1000231,	PLACE1000258,	PLACE1000442,	PLACE1000560,	PLACE1000912,	PLACE1000927,
	PLACE1001016,	PLACE1001100,	PLACE1001114,	PLACE1001183,	PLACE1001229,	PLACE1001340,
	PLACE1001407,	PLACE1001500,	PLACE1001516,	PLACE1001655,	PLACE1001836,	PLACE1001918,
35	PLACE1002080,	PLACE1002095,	PLACE1002153,	PLACE1002329,	PLACE1002374,	PLACE1002518,
	PLACE1002547,	PLACE1002726,	PLACE1002905,	PLACE1002911,	PLACE1002967,	PLACE1003163,
	PLACE1003407,	PLACE1003428,	PLACE1003438,	PLACE1003460,	PLACE1003529,	PLACE1003598,
	PLACE1003644,	PLACE1003772,	PLACE1003839,	PLACE1003845,	PLACE1003852,	PLACE1004078,
	PLACE1004166,	PLACE1004168,	PLACE1004199,	PLACE1004279,	PLACE1004282,	PLACE1004305,
40	PLACE1004441,	PLACE1004482,	PLACE1004492,	PLACE1004520,	PLACE1004630,	PLACE1004637,
	PLACE1004648,	PLACE1004816,	PLACE1004887,	PLACE1005005,	PLACE1005031,	PLACE1005383,
	PLACE1005410,	PLACE1005426,	PLACE1005539,	PLACE1005544,	PLACE1005569,	PLACE1005725,
	PLACE1005736,	PLACE1005768,	PLACE1005815,	PLACE1005878,	PLACE1005927,	PLACE1006071,
	PLACE1006073,	PLACE1006079,	PLACE1006277,	PLACE1006443,	PLACE1006716,	PLACE1006809,
45	PLACE1007077,	PLACE1007096,	PLACE1007626,	PLACE1007702,	PLACE1008469,	PLACE1008985,
	PLACE1009067,	PLACE1009527,	PLACE1009982,	PLACE1010078,	PLACE1010251,	PLACE1010445,
	PLACE1011045,	PLACE1011116,	PLACE1011181,	PLACE1011236,	PLACE1011364,	PLACE1011516,
	PLACE1011708,	PLACE1011978,	PLACE2000118,	PLACE2000219,	PLACE3000181,	PLACE4000354,
	PLACE4000455,	SKNMC1000014,	THYRO1000061,	THYRO1000099,	THYRO1000584,	THYRO1000795,
50	THYRO1000866,	THYRO1000999,	THYRO1001063,	THYRO1001113,	THYRO1001128,	THYRO1001205,
	THYRO1001237,	THYRO1001242,	THYRO1001456,	THYRO1001457,	THYRO1001478,	THYRO1001495,
	THYRO1001523,	THYRO1001529,	THYRO1001593,	THYRO1001608,	THYRO1001700,	THYRO1001702,
	THYRO1001725,	THYRO1001770,	THYRO1001803,	Y79AA1000127,	Y79AA1000207,	Y79AA1000226,
	Y79AA1000270,	Y79AA1000426,	Y79AA1000521,	Y79AA1000776,	Y79AA1000777,	Y79AA1000888,
55	Y79AA1000967,	Y79AA1001013,	Y79AA1001090,	Y79AA1001272,	Y79AA1001328,	Y79AA1001426,
	Y79AA1001427,	Y79AA1001430,	Y79AA1001523,	Y79AA1001530,	Y79AA1001592,	Y79AA1001727,
	Y79AA1001787,	Y79AA1001793,	Y79AA1001799,	Y79AA1001803,	Y79AA1001863,	Y79AA1002022,
	Y79AA1002213,	Y79AA1002373,	Y79AA1002376,	Y79AA1002381,		

EP 1 130 094 A2

[0207] Genes that were expressed in all the tissues tested are indicated below by the corresponding clone names:
clone: BNGH41000020, HEMBA1000300, HEMBA1001390, HEMBA1002239, HEMBA1002316, HEMBA1004007,
HEMBA1004067, HEMBA1005145, HEMBA1005230, HEMBA1005929, HEMBA1006357, HEMBA1006482,
HEMBA1006567, HEMBA1001847, HEMBA1001978, MAMMA1000614, MAMMA1000652, MAMMA1000810,
5 MAMMA1000814, MAMMA1001066, MAMMA1001094, MAMMA1001284, MAMMA1001310, MAMMA1001634,
MAMMA1002165, MAMMA1002205, MAMMA1002224, NT2RM1000462, NT2RM1000855, NT2RM1000858,
NT2RM2000423, NT2RM4000761, NT2RM4000997, NT2RP1000271, NT2RP1000325, NT2RP1000465,
NT2RP2001538, NT2RP2001662, NT2RP2001903, NT2RP2002015, NT2RP2002188, NT2RP2002409,
NT2RP2002510, NT2RP2002533, NT2RP2004556, NT2RP2004794, NT2RP2004847, NT2RP2005069,
10 NT2RP2005163, NT2RP2005535, NT2RP2006269, NT2RP3000171, NT2RP3000645, NT2RP3000838,
NT2RP3001271, NT2RP3001754, NT2RP3003076, NT2RP3003354, NT2RP3003614, NT2RP3004640,
NT2RP3004647, OVARC1000090, OVARC1000208, OVARC1000553, OVARC1000995, OVARC1001030,
OVARC1001049, PLACE1000231, PLACE1000258, PLACE1001516, PLACE1002080, PLACE1002911,
PLACE1003598, PLACE1004648, PLACE1006443, PLACE1008469, PLACE1011708, PLACE2000118,
15 THYRO1001128, THYRO1001205, THYRO1001242, THYRO1001803, Y79AA1000207, Y79AA1001013,
Y79AA1001272, Y79AA1001328, Y79AA1001793, Y79AA1001863, Y79AA1002022, Y79AA1002376.

[0208] Genes that were expressed at low levels in any of the tissues tested are indicated below by the corresponding
clone names: clone: HEMBA1000006, HEMBA1000128, HEMBA1000349, HEMBA1000590, HEMBA1000671,
HEMBA1000732, HEMBA1000745, HEMBA1000907, HEMBA1001184, HEMBA1001221, HEMBA1001272,
20 HEMBA1001297, HEMBA1001878, HEMBA1001886, HEMBA1002420, HEMBA1003497, HEMBA1003602,
HEMBA1003732, HEMBA1004250, HEMBA1004785, HEMBA1004971, HEMBA1004982, HEMBA1005267,
HEMBA1005522, HEMBA1005913, HEMBA1006299, HEMBA1006572, HEMBA1006724, HEMBA1007241,
HEMBA1000276, HEMBA1000407, HEMBA1000668, HEMBA1000679, HEMBA1000881, HEMBA1001200,
HEMBA1001547, HEMBA1002120, HEMBA1002245, MAMMA1000141, MAMMA1000496, MAMMA1001237,
25 MAMMA1001623, MAMMA1001978, MAMMA1002080, MAMMA1002087, MAMMA1002234, MAMMA1002633,
NT2RM1000580, NT2RM1000899, NT2RM2000632, NT2RM2001643, NT2PM2001818, NT2RM2001902,
NT2RM2001939, NT2RM4000100, NT2RM4000115, NT2RM4000284, NT2RM4000326, NT2RM4000417,
NT2RM4000587, NT2RM4000648, NT2RM4002352, NT2RP1000050, NT2RP1000239, NT2RP1000261,
NT2RP1000448, NT2RP1000551, NT2RP1000579, NT2RP1000613, NT2RP1000679, NT2RP1001004,
30 NT2RP1001020, NT2RP1001031, NT2RP1001563, NT2RP2000394, NT2RP2000514, NT2RP2000649,
NT2RP2000663, NT2RP2000712, NT2RP2000818, NT2RP2000903, NT2RP2001276, NT2RP2001495,
NT2RP2001755, NT2RP2001769, NT2RP2001817, NT2RP2001915, NT2RP2001948, NT2RP2002304,
NT2RP2002674, NT2RP2002976, NT2RP2003042, NT2RP2003545, NT2RP2003664, NT2RP2003931,
NT2RP2004495, NT2RP2004670, NT2RP2005514, NT2RP2005883, NT2RP2005994,
35 NT2RP2006134, NT2RP2006512, NT2RP3000169, NT2RP3000444, NT2RP3000481, NT2RP3000616,
NT2RP3000871, NT2RP3001012, NT2RP3001061, NT2RP3001240, NT2RP3001322, NT2RP3001542,
NT2RP3002286, NT2RP3002342, NT2RP3002448, NT2RP3002571, NT2RP3002664, NT2RP3002790,
NT2RP3002887, NT2RP3002983, NT2RP3003527, NT2RP3003535, NT2RP3003559, NT2RP3004000,
NT2RP3004075, NT2RP3004345, NT2RP4000962, NT2RP4001001, NT2RP4001009, NT2RP4001467,
40 NT2RP4002451, OVARC1000003, OVARC1000275, OVARC1000298, OVARC1000307, OVARC1000811,
OVARC1001132, OVARC1001222, OVARC1001338, OVARC1001607, OVARC1001725, OVARC1001727,
OVARC1002058, OVARC1002178, PLACE1000033, PLACE1000740, PLACE1000914, PLACE1000986,
PLACE1001123, PLACE1001231, PLACE1001401, PLACE1001464, PLACE1001536, PLACE1001564,
PLACE1001788, PLACE1001795, PLACE1001949, PLACE1002355, PLACE1003135, PLACE1003573,
45 PLACE1003737, PLACE1004028, PLACE1004450, PLACE1004519, PLACE1005003, PLACE1005239,
PLACE1005250, PLACE1005519, PLACE1005601, PLACE1005660, PLACE1005669, PLACE1005682,
PLACE1005745, PLACE1006093, PLACE1006208, PLACE1006219, PLACE1006290, PLACE1006515,
PLACE1006786, PLACE1006959, PLACE1007028, PLACE1007040, PLACE1007081, PLACE1007296,
PLACE1007591, PLACE1007845, PLACE1007881, PLACE1007971, PLACE1008282, PLACE1008297,
50 PLACE1008359, PLACE1008549, PLACE1008657, PLACE1008716, PLACE1008744, PLACE1008984,
PLACE1009196, PLACE1009279, PLACE1009546, PLACE1009600, PLACE1009735, PLACE1010011, PLACE1010081,
PLACE1010713, PLACE1010784, PLACE1010827, PLACE1010968, PLACE1011407, PLACE1011824,
PLACE3000213, SKNMC1000004, SKNMC1000082, THYRO1000036, THYRO1000196, THYRO1000400,
55 THYRO1000580, THYRO1000678, THYRO1000776, THYRO1000846, THYRO1000956, THYRO1001071,
THYRO1001102, THYRO1001266, THYRO1001327, THYRO1001471, Y79AA1000876, Y79AA1000959,
Y79AA1001056, Y79AA1001062, Y79AA1001264, Y79AA1001795.

[0209] Genes exhibiting characteristic features in the expression thereof were selected by statistical analysis of these

data. Two examples are shown below to describe the selection of genes of which expression is varied greatly among tissues. The β -actin gene is used frequently as a control in gene expression analysis. Genes of which expression is varied greatly among tissues as compared that of the β -actin gene were determined as follows. Specifically, sum of squared deviation was calculated in the signal intensity of β -actin observed in each tissue, which was divided by 7 degrees of freedom to determine a variance S_a^2 . Next, sum of squared deviation was calculated in the signal intensity of a compared gene in each tissue, which was divided by 7 degrees of freedom to determine a variance S_b^2 . By taking variance ratio F as $F=S_b^2/S_a^2$, genes with a significance level of 5% or more were extracted in the F distribution. Genes extracted are indicated below by the corresponding clone names: clone: BNGH41000020, NT2RM4000761, Y79AA1002376.

[0210] Gene of OVARC1000037(heterogeneous nuclear ribonucleoprotein (hnRNP)) which expression is varied little. Genes of which expression is varied greatly among tissues as compared that of the OVARC1000037 gene were determined as follows. Specifically, sum of squared deviation was calculated in the signal intensity of β -actin observed in each tissue, which was divided by 7 degrees of freedom to determine a variance S_a^2 . Next, sum of squared deviation was calculated in the signal intensity of a gene to be compared observed in each tissue, which was divided by 7 degrees of freedom to determine a variance S_b^2 . By taking variance ratio F as $F=S_b^2/S_a^2$, genes with a significance level of 5% or more were extracted in the F distribution. Genes extracted are indicated below by the corresponding clone names: clone: BNGH41000020, HEMBA1000300, OVARC1001030, NT2RM4000761, PLACE1000231, HEMBA1002316, NT2RP1000325, NT2RP1000271, PLACE1004648, HEMBA1005145, HEMBA1005929, NT2RP2002510, NT2RP2001538, NT2RP2002409, NT2RP2002188, NT2RP2001903, NT2RP2002533, NT2RP2002015, NT2RP2006269, NT2RP2004837, NT2RP2004205, NT2RP2005378, HEMBA1006357, HEMBB1000567, NT2RP2003940, NT2RP2004794, HEMBA1006912, NT2RP2004556, NT2RP2005163, NT2RP3000838, NT2RP3001271, PLACE2000118, NT2RP3000645, NT2RP3003076, HEMBB1002693, MAMMA1000046, NT2RP3003354, THYRO1001205, MAMMA1000614, MAMMA1000652, MAMMA1000810, THYRO1001242, MAMMA1001066, MAMMA1002224, MAMMA1001634, MAMMA1001094, MAMMA1002205, NT2RM1000855, NT2RM1000858, Y79AA1002376, NT2RM2000423.

[0211] Thus, characteristic features in the expression of a gene are illustrated by comparing and statistically analyzing the expression of many genes.

Analysis of disease-associated genes

[0212] Non-enzymic protein glycation reaction is believed to be a cause of a variety of chronic diabetic complications. Accordingly, genes of which expression is elevated or decreased in a glyated protein-specific manner in the endothelial cells are associated with diabetic complications caused by glyated proteins. Vascular endothelial cells are affected with glyated proteins present in blood. Reaction products of non-enzymic protein glycation include amadori compound (glycated protein) as a mildly glyated protein and advanced glycation endproduct as a heavily glyated protein. Hence, a survey was carried out for genes of which expression levels are varied depending on the presence of these glyated proteins in endothelial cells. The mRNAs were extracted from endothelial cells that were cultured in the presence or absence of glyated protein. The mRNAs were converted into radiolabeled first strand cDNAs for preparing probes. The probes were hybridized to the above-mentioned DNA array. Signal of each DNA spot was detected by BAS2000 and analyzed by ArrayGauge (Fuji Photo Film Co., Ltd.).

[0213] Advanced glycation endproduct of bovine serum albumin was prepared as follows: bovine serum albumin (BSA; Sigma) was incubated in a phosphate buffer solution containing 50 mM glucose at 37 for 8 weeks; and the resulting brownish BSA was dialyzed against a phosphate buffer solution.

[0214] Human normal pulmonary arterial endothelial cells (Cell Applications) were cultured in an Endothelial Cell Growth Medium (Cell Applications). The culture dish (Falcon) with the cells were incubated in a CO₂ incubator (37., 5% CO₂, in a humid atmosphere). When the cells were grown to be confluent in the dish, 250.g/ml of bovine serum albumin (sigma), glycated bovine serum albumin (Sigma) or advanced glycation endproduct of bovine serum albumin was added thereto and the cells were incubated for 33 hours. The mRNA was extracted from the cells by using a FastTack^(TM) 2.0 kit (Invitrogen). The labeling of hybridization probe was carried out by using the mRNA according to the same procedure as described above.

[0215] Table 185 shows the expression level of each cDNA in human pulmonary arterial endothelial cells cultured in a medium containing bovine serum albumin (sigma), glycated bovine serum albumin (Sigma) or advanced glycation endproduct of bovine serum albumin. Genes of which expression was detected in the endothelial cell are as follows: BNGH41000020, BNGH41000087, HEMBA1000275, HEMBA1000300, HEMBA1000477, HEMBA1000634, HEMBA1000671, HEMBA1000713, HEMBA1000745, HEMBA1000835, HEMBA1000875, HEMBA1000940, HEMBA1001390, HEMBA1002131, HEMBA1002163, HEMBA1002164, HEMBA1002195, HEMBA1002227, HEMBA1002239, HEMBA1002420, HEMBA1002767, HEMBA1002992, HEMBA1003047, HEMBA1003120, HEMBA1003294, HEMBA1003315, HEMBA1003602, HEMBA1003945, HEMBA1004007, HEMBA1004067,

EP 1 130 094 A2

	HEMBA1004971,	HEMBA1005145,	HEMBA1005267,	HEMBA1005337,	HEMBA1005698,	HEMBA1005929,
	HEMBA1005945,	HEMBA1006171,	HEMBA1006299,	HEMBA1006335,	HEMBA1006357,	HEMBA1006430,
	HEMBA1006482,	HEMBA1006658,	HEMBA1006724,	HEMBA1006770,	HEMBA1006912,	HEMBA1006960,
	HEMBA1007063,	HEMBB1000447,	HEMBB1000642,	HEMBB1000905,	HEMBB1001026,	HEMBB1001048,
5	HEMBB1001573,	HEMBB1001847,	HEMBB1001978,	HEMBB1002041,	HEMBB1002427,	HEMBB1002663,
	HEMBB1002693,	MAMMA1000102,	MAMMA1000106,	MAMMA1000204,	MAMMA1000403,	MAMMA1000449,
	MAMMA1000614,	MAMMA1000652,	MAMMA1000810,	MAMMA1000814,	MAMMA1000881,	MAMMA1000986,
	MAMMA1001066,	MAMMA1001237,	MAMMA1001284,	MAMMA1001344,	MAMMA1001615,	MAMMA1001634,
	MAMMA1001893,	MAMMA1001901,	MAMMA1001957,	MAMMA1002087,	MAMMA1002095,	MAMMA1002165,
10	MAMMA1002205,	MAMMA1002224,	MAMMA1002633,	MAMMA1003126,	NT2RM1000462,	NT2RM1000580,
	NT2RM1000789,	NT2RM1000855,	NT2RM1000858,			
	NT2RM2000241,	NT2RM2000306,	NT2RM2000410,	NT2RM2000423,	NT2RM2000582,	NT2RM2000589,
	NT2RM2000622,	NT2RM2000773,	NT2RM4000100,	NT2RM4000198,	NT2RM4000284,	NT2RM4000444,
	NT2RM4000587,	NT2RM4000593,	NT2RM4000761,	NT2RM4000997,	NT2RM4001321,	NT2RM4001325,
15	NT2RM4001377,	NT2RM4001735,	NT2RM4001768,	NT2RM4001843,	NT2RP1000002,	NT2RP1000181,
	NT2RP1000271,	NT2RP1000300,	NT2RP1000325,	NT2RP1000465,	NT2RP1000740,	NT2RP1000981,
	NT2RP2000092,	NT2RP2000240,	NT2RP2000479,	NT2RP2000533,	NT2RP2000610,	NT2RP2000616,
	NT2RP2000649,	NT2RP2000663,	NT2RP2000712,	NT2RP2000903,	NT2RP2001276,	NT2RP2001388,
	NT2RP2001480,	NT2RP2001495,	NT2RP2001529,	NT2RP2001538,	NT2RP2001662,	NT2RP2001878,
20	NT2RP2001903,	NT2RP2001948,	NT2RP2001956,	NT2RP2002015,	NT2RP2002188,	NT2RP2002232,
	NT2RP2002409,	NT2RP2002510,	NT2RP2002527,	NT2RP2002533,	NT2RP2002564,	NT2RP2002721,
	NT2RP2002824,	NT2RP2002942,	NT2RP2002976,	NT2RP2003138,	NT2RP2003210,	NT2RP2003390,
	NT2RP2003593,	NT2RP2003599,	NT2RP2003664,	NT2RP2003931,	NT2RP2003940,	NT2RP2004069,
	NT2RP2004108,	NT2RP2004179,	NT2RP2004205,	NT2RP2004495,	NT2RP2004524,	NT2RP2004556,
25	NT2RP2004606,	NT2RP2004648,	NT2RP2004794,	NT2RP2004837,	NT2RP2004847,	NT2RP2005027,
	NT2RP2005069,	NT2RP2005163,	NT2RP2005247,	NT2RP2005378,	NT2RP2005425,	NT2RP2005535,
	NT2RP2005541,	NT2RP2005632,	NT2RP2005774,			
	NT2RP2005878,	NT2RP2006099,	NT2RP2006134,	NT2RP2006269,	NT2RP2006512,	NT2RP3000011,
	NT2RP3000171,	NT2RP3000201,	NT2RP3000232,	NT2RP3000436,	NT2RP3000460,	NT2RP3000645,
30	NT2RP3000652,	NT2RP3000676,	NT2RP3000721,	NT2RP3000818,	NT2RP3000820,	NT2RP3000838,
	NT2RP3000907,	NT2RP3001159,	NT2RP3001195,	NT2RP3001240,	NT2RP3001271,	NT2RP3001388,
	NT2RP3001592,	NT2RP3001738,	NT2RP3001754,	NT2RP3002015,	NT2RP3002324,	NT2RP3002342,
	NT2RP3002353,	NT2RP3002409,	NT2RP3002448,	NT2RP3002721,	NT2RP3002737,	NT2RP3002738,
	NT2RP3002836,	NT2RP3002900,	NT2RP3003076,	NT2RP3003354,	NT2RP3003448,	NT2RP3003473,
35	NT2RP3003532,	NT2RP3003614,	NT2RP3003939,	NT2RP3003963,	NT2RP3004025,	NT2RP3004067,
	NT2RP3004075,	NT2RP3004083,	NT2RP3004090,	NT2RP3004119,	NT2RP3004130,	NT2RP3004133,
	NT2RP3004294,	NT2RP3004309,	NT2RP3004345,	NT2RP3004374,	NT2RP3004557,	NT2RP3004625,
	NT2RP3004640,	NT2RP3004647,	NT2RP4000108,	NT2RP4000634,	NT2RP4001001,	NT2RP4001009,
	NT2RP4001467,	NT2RP4001877,	NT2RP4001879,	NT2RP4002187,	NT2RP4002451,	NT2RP4002715,
40	OVARC1000003,	OVARC1000090,	OVARC1000105,	OVARC1000137,	OVARC1000208,	OVARC1000298,
	OVARC1000313,	OVARC1000331,	OVARC1000410,	OVARC1000439,	OVARC1000553,	OVARC1000775,
	OVARC1000853,	OVARC1000873,	OVARC1000916,	OVARC1000956,	OVARC1000995,	OVARC1001030,
	OVARC1001049,	OVARC1001086,	OVARC1001132,			
	OVARC1001222,	OVARC1001260,	OVARC1001336,	OVARC1001569,	OVARC1001570,	OVARC1001596,
45	OVARC1001607,	OVARC1001807,	OVARC1001991,	PLACE1000231,	PLACE1000258,	PLACE1000442,
	PLACE1000740,	PLACE1000927,	PLACE1001016,	PLACE1001100,	PLACE1001114,	PLACE1001123,
	PLACE1001229,	PLACE1001340,	PLACE1001407,	PLACE1001464,	PLACE1001788,	PLACE1001795,
	PLACE1001918,	PLACE1002080,	PLACE1002095,	PLACE1002329,	PLACE1002374,	PLACE1002518,
	PLACE1002547,	PLACE1002726,	PLACE1002905,	PLACE1002911,	PLACE1002967,	PLACE1003163,
50	PLACE1003407,	PLACE1003460,	PLACE1003573,	PLACE1003598,	PLACE1003644,	PLACE1003772,
	PLACE1003839,	PLACE1003845,	PLACE1004078,	PLACE1004166,	PLACE1004168,	PLACE1004199,
	PLACE1004279,	PLACE1004282,	PLACE1004441,	PLACE1004482,	PLACE1004492,	PLACE1004637,
	PLACE1004887,	PLACE1005003,	PLACE1005005,	PLACE1005031,	PLACE1005250,	PLACE1005410,
	PLACE1005519,	PLACE1005544,	PLACE1005660,	PLACE1005669,	PLACE1005725,	PLACE1005736,
55	PLACE1005745,	PLACE1005768,	PLACE1005815,	PLACE1006073,	PLACE1006208,	PLACE1006219,
	PLACE1006290,	PLACE1006443,	PLACE1006809,	PLACE1006959,	PLACE1007028,	PLACE1007296,
	PLACE1007626,	PLACE1007702,	PLACE1007845,	PLACE1008282,	PLACE1008469,	PLACE1008657,
	PLACE1009196,	PLACE1009600,	PLACE1003735,	PLACE1010081,	PLACE1010251,	PLACE1010713,

EP 1 130 094 A2

PLACE1011116, PLACE1011181,
 PLACE1011236, PLACE1011516, PLACE1011708, PLACE1011824, PLACE1011978, PLACE2000118,
 PLACE3000181, SKNMC1000004, SKNMC1000014, THYRO1000584, THYRO1000866, THYRO1001113,
 THYRO1001128, THYRO1001205, THYRO1001242, THYRO1001495, THYRO1001523, THYRO1001529,
 5 THYRO1001593, THYRO1001608, THYRO1001702, THYRO1001725, THYRO1001770, THYRO1001803,
 Y79AA1000117, Y79AA1000207, Y79AA1000226, Y79AA1000270, Y79AA1000426, Y79AA1000777,
 Y79AA1000876, Y79AA1000888, Y79AA1000959, Y79AA1001013, Y79AA1001056, Y79AA1001090,
 Y79AA1001264, Y79AA1001272, Y79AA1001328, Y79AA1001427, Y79AA1001430, Y79AA1001530,
 Y79AA1001592, Y79AA1001727, Y79AA1001793, Y79AA1001799, Y79AA1001863, Y79AA1002022,
 10 Y79AA1002213, Y79AA1002373, Y79AA1002376, Y79AA1002381.

[0216] Signal ratios of EC_AGE_BSA to EC_BSA and of EC_glycated_BSA to EC_BSA were calculated for each gene. Genes with high signal ratios were selected. In the case of calculating the ratio of signal value of 40 or less to that of more than 40, such signal values were, for convenience, taken as 40 instead of the real values. When the ratio EC_AGE_BSA/EC_BSA is 2 or more, expression of the genes exhibiting such ratio is expected to be elevated due to advanced glycation endproduct of bovine serum albumin. The higher the value is, the higher the gene expression level is. When the ratio EC_AGE_BSA/EC_BSA ranges from 0.5 to 2, expression of the genes exhibiting such ratio is expected to be unaffected due to advanced glycation endproduct of bovine serum albumin. When the ratio EC_AGE_BSA/EC_BSA is less than 0.5, expression of the genes exhibiting such ratio value is expected to be decreased due to advanced glycation endproduct of bovine serum albumin. The lower the value is, the lower the gene expression level is.

20 [0217] Clone with EC_AGE_BSA/BC_BSA ratio of 2 or higher are as follows: NT2RP2001538, NT2RP4001001 and Y79AA1000967.

[0218] These cDNAs are associated with diabetes.

Analysis of genes associated with neural cell differentiation

25 [0219] Genes involved in neural cell differentiation are useful for treating neurological diseases. It is possible that genes with varying expression levels in response to induction of cellular differentiation in neural cells are associated with neurological diseases.

30 [0220] A survey was performed for genes of which expression levels are varied in response to induction of differentiation (stimulation by retinoic acid (RA)) in cultured cells of a neural strain, NT2.

[0221] The NT2 cells were treated basically according to supplier's instruction manual. "Undifferentiated NT2 cells" means NT2 cells successively cultured in an Opti-MEM I (GIBCO-BRL; catalog No. 31985) containing 10%(v/v) fetal bovine serum and 1%(v/v) penicillin-streptomycin (GIBCO BRL). "NT2 cells cultured in the presence of retinoic acid" means the cells resulted from transferring undifferentiated NT2 cells into a retinoic acid-containing medium, which consists of D-MEM (GIBCO BRL; catalog No. 11965), 10%(v/v) fetal bovine serum, 1%(v/v) penicillin-streptomycin and 10.M retinoic acid (GIBCO-BRL), and the subsequent successive culture therein for 5 weeks. "NT2 cells that were cultured in the presence of retinoic acid and then further cultured in the presence of cell-division inhibitor added" means NT2 cells resulted from transferring NT2 cells cultured in the presence of retinoic acid for 5 weeks into a cell-division inhibitor-containing medium, which consisted of D-MEM(GIBCO BRL; catalog No.11965), 10%(v/v) fetal bovine serum, 1%(v/v) penicillin-streptomycin, 10. M retinoic acid, 10.M FudR (5-fluoro-2'-deoxyuridine: GIBCO BRL), 10. M Urd (Uridine: GIBCO BRL) and 1.M araC (Cytosine.-D-Arabinofuranoside: GIBCO BRL), and the subsequence successive culture for 2 weeks. Each of the cells were treated with trypsin and then harvested. Total RNAs were extracted from the cells by using S.N.A.P.(TM) Total RNA Isolation kit (Invitrogen). The labeling of probe used for hybridization was carried out by using 10.g of the total RNA according to the same methods as described above. The data were obtained in triplicate (n=3). The data of signal value representing gene expression level in the cells in the presence of stimulation for inducing differentiation were compared with those in the absence of the stimulation. The comparison was performed by statistical treatment of two-sample t-test. Clones with significant difference in the signal distribution were selected under the condition of p<0.05. In this analysis, clones with the difference can be statistically detected even when the signals were low. Accordingly, clones with signal value of 40 or less were also assessed for the selection.

50 [0222] Tables 186-365 show the expression level of each cDNA in undifferentiated NT2 cells, NT2 cells cultured in the presence of RA, and NT2 cells that were cultured in the presence of RA and that were further cultured in the presence of cell-division inhibitor added.

[0223] Averaged signal values (M_1 , M_2) and sample variances (s_1^2 , s_2^2) were calculated for each gene in each of the cells, and then, the pooled sample variances s^2 were obtained from the sample variances of the two types of cells to be compared. The t values were determined according to the following formula: $t=(M_1-M_2)/s/(1/3+1/3)^{1/2}$. When the determined t-value was greater than a t-value at P, which means the probability of significance level, of 0.05 or 0.01 in the t-distribution table with 4 degrees of freedom, the difference was judged to be found in the expression level of the gene between the two types of cells at p<0.05 or p<0.01, respectively. The tables also include the information on

EP 1 130 094 A2

an increase (+) or decrease (-) in the expression level of a gene in the treated cells when the level is compared with that of untreated undifferentiated cells.

[0224] Clones of which expression levels increased by RA are as follows: HEMBA1000121, HEMBA1000275,

5 HEMBA1000300, HEMBA1000634, HEMBA1000671, HEMBA1000875, HEMBA1001184, HEMBA1001390,
HEMBA1001886, HEMBA1002163, HEMBA1002227, HEMBA1002420, HEMBA1002421, HEMBA1003072,
HEMBA1003120, HEMBA1003294, HEMBA1003497, HEMBA1004007, HEMBA1004110, HEMBA1004391,
HEMBA1004444, HEMBA1005230, HEMBA1005246, HEMBA1005267, HEMBA1005489, HEMBA1005913,
HEMBA1006299, HEMBA1006357, HEMBA1006517, HEMBA1006544, HEMBA1006658, HEMBA1006749,
HEMBA1007063, HEMBA1007241, HEMBB1000447, HEMBB1000542, HEMBB1000567, HEMBB1000642,
10 HEMBB1000668, HEMBB1001026, HEMBB1001847, HEMBB1002051, HEMBB1002120, HEMBB1002228,
HEMBB1002693, MAMMA1000106, MAMMA1000141, MAMMA1000473, MAMMA1000528, MAMMA1000810,
MAMMA1000881, MAMMA1001634, MAMMA1001957, MAMMA1002205, MAMMA1002224, NT2RM2000423,
NT2RM2000497, NT2RM2000582, NT2RM2001126, NT2RM2001902, NT2RM4000198, NT2RM4000284,
NT2RM4000593, NT2RM4001321, NT2RP1000002, NT2RP1000050, NT2RP1000181, NT2RP1000261,
15 NT2RP1000465, NT2RP1000468, NT2RP1000579, NT2RP1000679, NT2RP2000092, NT2RP2000479,
NT2RP2000610, NT2RP2000663, NT2RP2000694, NT2RP2000903, NT2RP2001388, NT2RP2001538,
NT2RP2001878, NT2RP2001015, NT2RP2002304, NT2RP2002721, NT2RP2002824, NT2RP2002942,
NT2RP2002974, NT2RP2002976, NT2RP2003179, NT2RP2003302, NT2RP2003383, NT2RP2003469,
NT2RP2003664, NT2RP2003940, NT2RP2004069, NT2RP2004108, NT2RP2004524, NT2RP2004556,
20 NT2RP2004670, NT2RP2005069, NT2RP2005247, NT2RP2005425, NT2RP2005463, NT2RP2005514,
NT2RP2005535, NT2RP2005541, NT2RP2005774, NT2RP2005878, NT2RP2005883, NT2RP2005887,
NT2RP2006099, NT2RP2006134, NT2RP3000011, NT2RP3000125, NT2RP3000171, NT2RP3000232,
NT2RP3000460, NT2RP3000481, NT2RP3000652, NT2RP3000677, NT2RP3000818, NT2RP3000820,
NT2RP3001044, NT2RP3001061, NT2RP3001170, NT2RP3001240, NT2RP3001322, NT2RP3001388,
25 NT2RP3001542, NT2RP3001592, NT2RP3001976, NT2RP3002790, NT2RP3002900, NT2RP3002983,
NT2RP3003000, NT2RP3003354, NT2RP3003532, NT2RP3003729, NT2RP3003874, NT2RP3003939,
NT2RP3004025, NT2RP3004083, NT2RP3004090, NT2RP3004130, NT2RP3004202, NT2RP3004294,
NT2RP3004640, NT2RP4000108, NT2RP4000634, NT2RP4002451, NT2RP4002715, OVARC1000090,
OVARC1000208, OVARC1000275, OVARC1000553, OVARC1000775, OVARC1000853, OVARC1000873,
30 OVARC1000916, OVARC1000995, OVARC1001030, OVARC1001049, OVARC1001132, OVARC1001596,
OVARC1002178, PLACE1000258, PLACE1000442, PLACE1000927, PLACE1000986, PLACE1001100,
PLACE1001123, PLACE1001795, PLACE1002518, PLACE1002547, PLACE1002967, PLACE1003407,
PLACE1003428, PLACE1003644, PLACE1003839, PLACE1004078, PLACE1004441, PLACE1004450,
PLACE1005669, PLACE1005682, PLACE1005736, PLACE1005768, PLACE1005815, PLACE1006073,
35 PLACE1006208, PLACE1007296, PLACE1007626, PLACE1008282, PLACE1008984, PLACE1008985,
PLACE1010445, PLACE1011708, PLACE1011978, PLACE4000455, SKNMC1000004, THYRO1000036,
THYRO1000580, THYRO1000776, THYRO1000999, THYRO1001063, THYRO1001128, THYRO1001205,
THYRO1001327, THYRO1001523, THYRO1001725, THYRO1001770, Y79AA1000207, Y79AA1000226,
Y79AA1000270, Y79AA1001056, Y79AA1001062, Y79AA1001090, Y79AA1001727, Y79AA1002213,
40 Y79AA1002381.

[0225] Clones of which expression levels decreased by RA are as follows: BNGH41000020, HEMBA1005070, NT2RP2005027, NT2RP3003473, Y79AA1002376.

[0226] Clones of which expression levels increase by RA/inhibitor are as follows:

45 HEMBA1000128, HEMBA1000875, HEMBA1001390, HEMBA1002163, HEMBA1002227, HEMBA1002421,
HEMBA1004391, HEMBA1004454, HEMBA1004785, HEMBA1005913, HEMBA1006171, HEMBA1006299,
HEMBA1006335, HEMBA1006544, HEMBA1007241, HEMBB1000447, HEMBB1000668, MAMMA1000994,
MAMMA1001344, NT2RM2000582, NT2RP1001004, NT2RP2000663, NT2RP2000694, NT2RP2000903,
NT2RP2001388, NT2RP2002674, NT2RP2002974, NT2RP2003383, NT2RP2004069, NT2RP2004606,
NT2RP2004837, NT2RP2005069, NT2RP2005425, NT2RP2005463, NT2RP2005541, NT2RP2005883,
50 NT2RP2005887, NT2RP3000460, NT2RP3000838, NT2RP3001044, NT2RP3001240, NT2RP3001388,
NT2RP3002721, NT2RP3002738, NT2RP3003469, NT2RP3004083, NT2RP3004130, NT2RP3004202,
NT2RP3004294, NT2RP3004640, NT2RP4000108, NT2RP4002451, NT2RP4002715, OVARC1000275,
OVARC1000467, OVARC1000553, OVARC1000853, OVARC1000873, OVARC1000916, OVARC1000995,
OVARC1001030, OVARC1001222, OVARC1001596, OVARC1002058, OVARC1002178, PLACE1000927,
55 PLACE1001123, PLACE1001407, PLACE1001464, PLACE1001564, PLACE1001795, PLACE1002547,
PLACE1003407, PLACE1003644, PLACE1003845, PLACE1004441, PLACE1004482, PLACE1005410,
PLACE1005601, PLACE1005725, PLACE1005736, PLACE1006093, PLACE1006219, PLACE1006290,
PLACE1006716, PLACE1007296, PLACE1007626, PLACE1008359, PLACE1010968, PLACE1011364,

EP 1 130 094 A2

PLACE1011824, THYRO1000678, THYRO1000776, THYRO1000999, THYRO1001113, THYRO1001237, THYRO1001523, Y79AA1000226, Y79AA1000888, Y79AA1001430.

[0227] Clones of which expression levels decrease by RA/inhibitor are as follows: HEMBA1000349, HEMBA1001297, HEMBA1001878, HEMBA1005070, HEMBA1006482, HEMBB1001959, NT2RM2001939, NT2RP1000981, NT2RP2001469, NT2RP3003473, OVARC1001132, PLACE1001655, Y79AA1000127, Y79AA1002381.

[0228] Clones of which expression levels increase in the presence of both RA and RA/inhibitor are as follows: HEMBA1000875, HEMBA1001390, HEMBA1002163, HEMBA1002227, HEMBA1002421, HEMBA1004391, HEMBA1005913, HEMBA1006299, HEMBA1006544, HEMBA1007241, HEMBB1000447, HEMBB1000668, NT2RM2000582, NT2RP2000663, NT2RP2000694, NT2RP2000903, NT2RP2001388, NT2RP2002974, NT2RP2003383, NT2RP2004069, NT2RP2005069, NT2RP2005425, NT2RP2005463, NT2RP2005541, NT2RP2005883, NT2RP2005887, NT2RP3000460, NT2RP3001044, NT2RP3001240, NT2RP3001388, N12RP30004083, NT2RP3004130, NT2RP3004202, NT2RP3004294, NT2RP3004640, NT2RP4000108, NT2RP4002451, NT2RP4002715, OVARC1000275, OVARC1000553, OVARC1000853, OVARC1000873, OVARC1000916, OVARC1000995, OVARC1001030, OVARC1001596, OVARC1002178, PLACE1000927, PLACE1001123, PLACE1001795, PLACE1002547, PLACE1003407, PLACE1003644, PLACE1004441, PLACE1005736, PLACE1007296, PLACE1007626, THYRO1000776, THYRO1000999, THYRO1001523, Y79AA1000226.

[0229] Clones of which expression levels decrease in the presence of both RA and RA/inhibitor are as follows: HEMBA1005070 and NT2RP3003473.

[0230] These are neurological disease-associated clones.

Analysis of rheumatoid arthritis-associated genes

[0231] The onset of rheumatoid arthritis is thought to be involved in the proliferation of synovial cells covering inner surfaces of joint cavity and in inflammatory reaction resulted from the action of cytokines produced by leukocytes infiltrating into the joint synovial tissues (Rheumatism Information Center.<http://www.rheuma-net.or.jp/>). Recent studies have also revealed that tissue necrosis factor (TNF)- α participates in the onset (Current opinion in immunology 1999, 11, 657-662). When the expression of a gene exhibits responsiveness to the action of TNF on synovial cells, the gene is considered to be involved in rheumatoid arthritis.

[0232] A survey was performed for genes of which expression levels are varied in response to TNF- α in the primary cell culture of synovial tissue. The primary cultured cells of the smooth muscle (Cell Applications) were grown to be confluent in a culture dish, and then, human TNF- α (Boehringer-Mannheim) was added at a final concentration of 10 ng/ml thereto. The culture was further continued for 24 hours.

[0233] Total RNA was extracted from the cells by using S.N.A.P.^(TM) Total RNA Isolation kit (Invitrogen). The labeling of probe used for hybridization was carried out by using 10 g of the total RNA according to the same methods as described above. The data were obtained in triplicate ($n=3$). The data of signal value representing gene expression level in the cells in the presence of TNF stimulation were compared with those in the absence of the stimulation. The comparison was performed by statistical treatment of two-sample t-test. Clones with significant difference in the signal distribution were selected under the condition of $p<0.05$. In this analysis, clones with the difference can be statistically detected even when the signals were low. Accordingly, clones with signal value of 40 or less were also assessed for the selection.

[0234] Table 366 shows the expression level of each cDNA in synovial cells cultured in the absence or presence of TNF.

[0235] Averaged signal values (M_1 , M_2) and sample variances (s_1^2 , s_2^2) for each gene were calculated in each of the cells, and then, the pooled sample variances s^2 were obtained from the sample variances of the two types of cells to be compared. The t-values were determined according to the following formula: $t=(M_1-M_2)/s/(1/3+1/3)^{1/2}$. When the determined t-value was greater than a t-value at P, which means the probability of significance level, of 0.05 or 0.01 in the t-distribution table with 4 degrees of freedom, the difference was judged to be found in the expression level of the gene between the two types of cells at $p<0.05$ or $p<0.01$, respectively. The tables also include the information of an increase (+) or decrease (-) in the expression level of a gene in the stimulated cells when the level is compared with that of unstimulated cells.

[0236] Clones of which expression levels are elevated by TNF- α are as follows:

BNGH41000020, HEMBA1000349, HEMBA1000634, HEMBA1000671, HEMBA1000835, HEMBA1000962, HEMBA1002178, HEMBA1002195, HEMBA1002239, HEMBA1002420, HEMBA1002524, HEMBA1002992, HEMBA1003315, HEMBA1003392, HEMBA1003487, HEMBA1003602, HEMBA1004067, HEMBA1004797, HEMBA1005337, HEMBA1005489, HEMBA1006916, HEMBB1000668, HEMBB1000905, HEMBB1001547, HEMBB1001573, HEMBB1002041, HEMBB1002663, MAMMA1000652, MAMMA1000810, MAMMA1001634,

EP 1 130 094 A2

MAMMA1002091, MAMMA1002234, NT2RM2000306, NT2RM4000417, NT2RP1000002, NT2RP1000181,
 NT2RP1000740, NT2RP2000694, NT2RP2001921, NT2RP2002527, NT2RP2004495, NT2RP2004606,
 NT2RP2005163, NT2RP2005463, NT2RP2006134, NT2RP3000171, NT2RP3000652, NT2RP3001195,
 NT2RP3001976, NT2RP3003473, NT2RP3003874, NT2RP3004090, NT2RP3004294, NT2RP3004557,
 5 NT2RP3004647, NT2RP4000108, NT2RP4001001, NT2RP4001877, OVARC1000090, OVARC1000105,
 OVARC1000275, OVARC1000439, OVARC1001607, PLACE1000740, PLACE1000927, PLACE1001016,
 PLACE1001100, PLACE1001464, PLACE1001500, PLACE1001918, PLACE1002095, PLACE1002547,
 PLACE1003644, PLACE1004519, PLACE1005031, PLACE1005410, PLACE1005736, PLACE1006219,
 PLACE1006809, PLACE1008716, PLACE1010081, THYRO1001770, Y79AA1000127, Y79AA1000207,
 10 Y79AA1000270, Y79AA1000876, Y79AA1001013, Y79AA1001264, Y79AA1001272, Y79AA1001328,
 Y79AA1001430, Y79AA1001530, Y79AA1001799.
[0237] Clones of which expression levels decrease by TNF- α are as follows:
 NT2RM4000326, NT2RP1000300, NT2RP2000514, NT2RP2001755, NT2RP2006042, NT2RP3000481,
 NT2RP3002790. These are rheumatoid arthritis-associated clones.

EXAMPLE 16

Search for a signal sequence, transmembrane region and functional domain in deduced amino acid sequences

20 **[0238]** The deduced amino acid sequences from the full-length nucleotide sequences were examined to predict the
 presence of a signal sequence in their amino-termini as well as the presence of a transmembrane region. The amino
 acid sequences were also searched for a protein functional domain (motif). The examinations for a signal sequence
 in the amino-terminus, for a transmembrane region and for a functional domain were performed by using PSORT [K.
 Nakai & M. Kanehisa, Genomics, 14:897-911 (1992)], SOSUI [T. Hirokawa et al., Bioinformatics, 14:378-379 (1998)]
 25 (Mitsui Knowledge Industry Co., Ltd.) and Pfam (<http://www.sanger.ac.uk/Software/Pfam/index.shtml>), respectively.
 When the presence of a signal sequence or a transmembrane region in the amino-terminus was predicted in the amino
 acid sequence by PSORT or SOSUI, the protein was predicted to be a secretory protein or a transmembrane protein.
 When the amino acid sequence matched a functional domain in the Pfam search for a functional domain, the function
 of the protein is predictable based on the matching data, for example, by referring to the functional categories in
 30 PROSITE (<http://www.expasy.ch/cgi-bin/prosite-list.pl>). The functional domain search can be performed by using
 PROSITE instead of Pfam.

[0239] Search results obtained by using the respective software programs are indicated below.

[0240] Clones whose deduced amino acid sequences were predicted to have signal sequences by PSORT search
 are as follows:
 35 HEMBA1000713, HEMBA1002420, HEMBA1002421, HEMBA1003101, HEMBA1004110, HEMBA1006707,
 HEMBA1006902, HEMBB1001530, HEMBB1001573, HEMBB1001978, HEMBB1002162, HEMBB1002245,
 HEMBB1002427, MAMMA1000102, MAMMA1000118, MAMMA1000457, MAMMA1001043, MAMMA1001344,
 MAMMA1001893, MAMMA1002070, MAMMA1002165, MAMMA1002633, NT2RM2000241, NT2RM2000410,
 NT2RM2001941, NT2RM4001325, NT2RP1001563, NT2RP2001495, NT2RP2002063, NT2RP2002721,
 40 NT2RP2003383, NT2RP2003593, NT2RP2003655, NT2RP2003664, NT2RP2004179, NT2RP2004205,
 NT2RP2004524, NT2RP2005463, NT2RP3000460, NT2RP3001012, NT2RP3001858, NT2RP3002836,
 NT2RP3003076, NT2RP3003532, NT2RP3004133, NT2RP3004309, NT2RP4001467, NT2RP4002451,
 OVARC1000298, OVARC1000811, PLACE1000231, PLACE1000740, PLACE1001183, PLACE1001536,
 PLACE1001564, PLACE1002095, PLACE1002374, PLACE1003839, PLACE1001482, PLACE1005005,
 45 PLACE1005250, PLACE1005383, PLACE1005410, PLACE1005544, PLACE1005569, PLACE1006093,
 PLACE1006277, PLACE1006809, PLACE1007626, PLACE1008359, PLACE1009067, PLACE1010251,
 PLACE1011236, SKNMC1000004, SKNMC1000014, THYRO1000099, THYRO1000196, THYRO1001237,
 THYRO1001327, THYRO1001523, THYRO1001702, THYRO1001725, Y79AA1000426, Y79AA1000521,
 Y79AA1000959, Y79AA1001013, Y79AA1001264, Y79AA1001328, Y79AA1001427, Y79AA1001430,
 50 Y79AA1001795, Y79AA1001803, Y79AA1002022.

[0241] Clones whose deduced amino acid sequences were predicted to have transmembrane regions by SOSUI
 search are as follows: BNGH41000091, HEMBA1000121, HEMBA1000349, HEMBA1000477, HEMBA1000713,
 HEMBA1000940, HEMBA1000962, HEMBA1001221, HEMBA1001228, HEMBA1001621, HEMBA1002167,
 HEMBA1002195, HEMBA1002227, HEMBA1002421, HEMBA1003101, HEMBA1003392, HEMBA1003530,
 55 HEMBA1003732, HEMBA1003945, HEMBA1004391, HEMBA1004454, HEMBA1004797, HEMBA1004982,
 HEMBA1005449, HEMBA1005522, HEMBA1005545, HEMBA1005698, HEMBA1006171, HEMBA1006299,
 HEMBA1006311, HEMBA1006335, HEMBA1006357, HEMBA1006430, HEMBA1006724, HEMBA1006960,
 HEMBB1000407, HEMBB1000447, HEMBB1000567, HEMBB1000679, HEMBB1000905, HEMBB1001026,

EP 1 130 094 A2

HEMBB1001407, HEMBB1001573, HEMBB1001978, HEMBB1002041, HEMBB1002162, HEMBB1002245,
 HEMBB1002427, HEMBB1002693, MAMMA1000102, MAMMA1000106, MAMMA1000118, MAMMA1000141,
 MAMMA1000204, MAMMA1000226, MAMMA1000457, MAMMA1000473, MAMMA1000591, MAMMA1000681,
 MAMMA1000810, MAMMA1000986, MAMMA1001043, MAMMA1001141, MAMMA1001237, MAMMA1001344,
 5 MAMMA1001893, MAMMA1001957, MAMMA1001978, MAMMA1002070, MAMMA1002091, MAMMA1002095,
 MAMMA1002633, NT2RM1000580, NT2RM1000855, NT2RM1000858, NT2RM2000410, NT2RM2000565,
 NT2RM2001626, NT2RM2001939, NT2RM2001941, NT2RM4000444, NT2RM4000587, NT2RM4000648,
 NT2RM4000997, NT2RM4001325, NT2RM4001735, NT2RM4001768, NT2RM4002352, NT2RP1000050,
 NT2RP1000181, NT2RP1000261, NT2RP1000300, NT2RP1000448, NT2RP1000551, NT2RP1000613,
 10 NT2RP1000981, NT2RP1001563, NT2RP2000479, NT2RP2000533, NT2RP2000649,
 NT2RP2000663, NT2RP2000694, NT2RP2000818, NT2RP2000903, NT2RP2001200, NT2RP2001276,
 NT2RP2001495, NT2RP2001915, NT2RP2001956, NT2RP2002188, NT2RP2002232, NT2RP2002527,
 NT2RP2002533, NT2RP2002721, NT2RP2002824, NT2RP2002942, NT2RP2002976, NT2RP2003042,
 NT2RP2003390, NT2RP2003469, NT2RP2003593, NT2RP2003655, NT2RP2003664, NT2RP2003950,
 15 NT2RP2004179, NT2RP2004205, NT2RP2004495, NT2RP2004524, NT2RP2004556, NT2RP2004606,
 NT2RP2004648, NT2RP2004794, NT2RP2005163, NT2RP2005181, NT2RP2005463, NT2RP2005597,
 NT2RP2005666, NT2RP2005883, NT2RP2005994, NT2RP2006004, NT2RP2006269, NT2RP2006512,
 NT2RP2006580, NT2RP3000169, NT2RP3000171, NT2RP3000304, NT2RP3000460, NT2RP3000616,
 NT2RP3000721, NT2RP3000818, NT2RP3000907, NT2RP3000921, NT2RP3001159, NT2RP3001195,
 20 NT2RP3001240, NT2RP3001271, NT2RP3001322, NT2RP3001388, NT2RP3001560, NT2RP3001592,
 NT2RP3001650, NT2RP3001738, NT2RP3002015, NT2RP3002311, NT2RP3002342, NT2RP3002411,
 NT2RP3002790, NT2RP3002836, NT2RP3002900, NT2RP3002958, NT2RP3003000, NT2RP3003354,
 NT2RP3003532, NT2RP3003535, NT2RP3003614, NT2RP3004025, NT2RP3004075, NT2RP3004083,
 NT2RP3004090, NT2RP3004130, NT2RP3004294, NT2RP3004309, NT2RP3004345, NT2RP3004406,
 25 NT2RP3004481, NT2RP3004552, NT2RP4001001, NT2RP4001009, NT2RP4001467, NT2RP4001879,
 NT2RP4002187, NT2RP4002451, NT2RP4002750, OVARC1000003, OVARC1000105, OVARC1000307,
 OVARC1000439, OVARC1000553, OVARC1001030, OVARC1001336,
 OVARC1001570, PLACE1000231, PLACE1000560, PLACE1000740, PLACE1000912, PLACE1000914,
 PLACE1000927, PLACE1001016, PLACE1001183, PLACE1001231, PLACE1001401, PLACE1001407,
 30 PLACE1001464, PLACE1001536, PLACE1001564, PLACE1001655, PLACE1001836, PLACE1001918,
 PLACE1001949, PLACE1002518, PLACE1002726, PLACE1002967, PLACE1003573, PLACE1003737,
 PLACE1003839, PLACE1003845, PLACE1003852, PLACE1004279, PLACE1004282, PLACE1004441,
 PLACE1004637, PLACE1004648, PLACE1004816, PLACE1004887, PLACE1005003, PLACE1005005,
 PLACE1005410, PLACE1005544, PLACE1005569, PLACE1005660, PLACE1005725, PLACE1005745,
 35 PLACE1005927, PLACE1006290, PLACE1006443, PLACE1006959, PLACE1007096, PLACE1007296,
 PLACE1007626, PLACE1007881, PLACE1008359, PLACE1008469, PLACE1008716, PLACE1008985,
 PLACE1009196, PLACE1009279, PLACE1009527, PLACE1009546, PLACE1009600, PLACE1010011,
 PLACE1010078, PLACE1010445, PLACE1010713, PLACE1010784, PLACE1010968, PLACE1011236,
 PLACE1011516, PLACE3000181, THYRO1000400, THYRO1000678, THYRO1000776, THYRO1000956,
 40 THYRO1001102, THYRO1001113, THYRO1001205, THYRO1001237, THYRO1001242, THYRO1001266,
 THYRO1001327, THYRO1001478, THYRO1001523, THYRO1001641, THYRO1001702, THYRO1001725,
 Y79AA1000207, Y79AA1000226, Y79AA1000270, Y79AA1000521, Y79AA1000888, Y79AA1001013,
 Y79AA1001212, Y79AA1001264, Y79AA1001328, Y79AA1001426, Y79AA1001427, Y79AA1001727,
 Y79AA1001787, Y79AA1001795, Y79AA1001803, Y79AA1002058,
 45 Y79AA1002129, Y79AA1002213, Y79AA1002373,

[0242] Names of clones whose deduced amino acid sequences were predicted to have functional domains by Pfam search, and names of the matched functional domains are shown below. When multiple functional domains matched a clone, each domain name was indicated, separated by a double-slash mark//.

50 HEMBA1000006//Src homology domain 3
 HEMBA1000128//SCP-like extracellular Proteins
 HEMBA1000349//ABC transporters
 HEMBA1000462//RNA recognition motif. (aka RRM, RBD, or RNP domain)
 HEMBA1000590//EGF-like domain//von Willebrand factor type A domain
 55 HEMBA1000671//Zinc finger, C2H2 type
 HEMBA1000732//EGF-like domain
 HEMBA1000940//Connexin
 HEMBA1001221//EGF-like domain//Kazal-type serine protease inhibitor domain

EP 1 130 094 A2

HEMBA1001621//7 transmembrane receptor (rhodopsin family)
 HEMBA1001878//WD domain, G-beta repeats
 HEMBA1002048//Zinc finger, C2H2 type
 HEMBA1002167//Carboxylesterases
 5 HEMBA1002551//WD domain, G-beta repeats
 HEMBA1002992//Ubiquitin family
 HEMBA1003047//CUB domain
 HEMBA1003120//Zinc finger, C2H2 type
 HEMBA1003230//EGF-like domain
 10 HEMBA1003392//Low-density lipoprotein receptor domain class A
 HEMBA1003497//Zinc finger, C2H2 type
 HEMBA1004250//Cadherin
 HEMBA1004391//Fibronectin type III domain//IG superfamily
 HEMBA1004454//4 transmembrane segments integral membrane proteins
 15 HEMBA1004785//'chromo' (CHRromatin Organization MODifier) domain
 HEMBA1005246//Zinc finger, C2H2 type
 HEMBA1005267//Ank repeat
 HEMBA1005545//7 transmembrane receptor (rhodopsin family)
 HEMBA1005929//Eukaryotic protein kinase domain
 20 HEMBA1005945//Mitochondrial carrier proteins
 HEMBA1006572//Zinc finger, C2H2 type
 HEMBA1006707//EGF-like domain//von Willebrand factor type A domain
 HEMBA1006749//EGF-like domain//von Willebrand factor type A domain
 HEMBA1006770//RNA recognition motif. (aka RRM, RBD, or RNP domain)
 25 HEMBA1006902//EGF-like domain//von Willebrand factor type A domain
 HEMBB1000106//Zinc finger, CCHC class
 HEMBB1000668//WD domain, G-beta repeats
 HEMBB1000881//Thrombospondin type 1 domain
 HEMBB1000905//WD domain, G-beta repeats
 30 HEMBB1002041//EGF-like domain//Kazal-type serine protease inhibitor domain
 HEMBB1002245//IG superfamily
 HEMBB1002302//Zinc finger, CCHC class
 HEMBB1002465//Acyl-CoA dehydrogenases
 HEMBB1002661//Helix-loop-helix DNA-binding domain
 35 MAMMA1000204//C2 domain
 MAMMA1000457//FAD/NAD-binding domain in oxidoreductases
 MAMMA1000681//7 transmembrane receptor (rhodopsin family)
 MAMMA1000881//Eukaryotic protein kinase domain//Protein kinase C terminal domain
 MAMMA1001150//Phorbol esters / diacylglycerol binding domain//Eukaryotic protein kinase domain
 40 MAMMA1001310//WD domain, G-beta repeats
 MAMMA1001532//Zinc finger, C2H2 type
 MAMMA1001615//Helix-loop-helix DNA-binding domain
 MAMMA1002070//Kringle domain
 MAMMA1002080//Ras family (contains ATP/GTP binding P-loop)
 45 MAMMA1002095//E1-E2 ATPases
 MAMMA1002165//Insulin-like growth factor binding proteins
 NT2RM1000789//HMG (high mobility group) box
 NT2RM1000855//eubacterial secY protein
 NT2RM1000899//Mitochondrial carrier proteins
 50 NT2RM2000589//PH (pleckstrin homology) domain
 NT2RM2000632//Helicases conserved C-terminal domain
 NT2RM2001792//Fibrinogen beta and gamma chains, C-terminal globular domain
 NT2RM2001902//Eukaryotic protein kinase domain
 NT2RM2001939//7 transmembrane receptor (rhodopsin family)
 55 NT2RM2001941//7 transmembrane receptor (rhodopsin family)
 NT2RM4000284//Class I Histocompatibility antigen, domains alpha 1 and 2
 NT2RM4000326//Zinc finger, C2H2 type
 NT2RM4000417//C2 domain

EP 1 130 094 A2

NT2RM4000444//ABC transporters
 NT2RM4001377//PH (pleckstrin homology) domain
 NT2RM4001768//Alcohol/other dehydrogenases, short chain type
 NT2RM4002352//Low-density lipoprotein receptor domain class A
 5 NT2RP1000181//Heme-binding domain in cytochrome b5 and oxidoreductases
 NT2RP1000271//Zinc finger, C2H2 type
 NT2RP1000325//Mitochondrial carrier proteins
 NT2RP1000613//Eukaryotic-type carbonic anhydrases
 NT2RP1000981//IG superfamily
 10 NT2RP1001004//Thrombospondin type 1 domain
 NT2RP1001020//Eukaryotic protein kinase domain
 NT2RP1001031//WD domain, G-beta repeats
 NT2RP1001563//EGF-like domain//Lectin C-type domain short and long forms//SCP-like extracellular Proteins
 NT2RP2000092//Zinc finger, C2H2 type
 15 NT2RP2000514//Fibronectin type III domain//IG superfamily
 NT2RP2000649//Zinc-binding metalloprotease domain
 NT2RP2000712//Zinc finger, C2H2 type
 NT2RP2000739//Zinc finger, C2H2 type
 NT2RP2001514//E1-E2 ATPases
 20 NT2RP2001529//Eukaryotic protein kinase domain
 NT2RP2001755//Thrombospondin type 1 domain
 NT2RP2001769//Eukaryotic protein kinase domain
 NT2RP2002188//Carboxylesterases
 NT2RP2002527//Heme-binding domain in cytochrome b5 and oxidoreductases
 25 NT2RP2002564//Zinc finger, C2H2 type
 NT2RP2002942//IG superfamily
 NT2RP2003179//Eukaryotic protein kinase domain
 NT2RP2003302//Zinc finger, C2H2 type
 NT2RP2003390//DnaJ, prokaryotic heat shock protein
 30 NT2RP2003469//Sugar (and other) transporters
 NT2RP2003545//Eukaryotic protein kinase domain
 NT2RP2003593//Thioredoxins
 NT2RP2003940//Zinc finger, C2H2 type
 NT2RP2004108//Zinc finger, C2H2 type
 35 NT2RP2004205//IG superfamily
 NT2RP2004670//Eukaryotic protein kinase domain
 NT2RP2004847//Zinc finger, C2H2 type
 NT2RP2005181//Amino acid permeases
 NT2RP2005247//Zinc finger, C3HC4 type (RING finger)
 40 NT2RP2005391//Fibronectin type III domain
 NT2RP2005535//Zinc finger, C2H2 type
 NT2RP2005774//Zinc finger, C2H2 type
 NT2RP2005878//Alcohol/other dehydrogenases, short chain type
 NT2RP2005941//Homeobox domain//'Paired box' domain
 45 NT2RP2006004//Fibronectin type III domain
 NT2RP3000011//WD domain, G-beta repeats
 NT2RP3000022//Eukaryotic protein kinase domain
 NT2RP3000063//Zinc finger, C2H2 type
 NT2RP3000148//Zinc finger, C2H2 type
 50 NT2RP3000172//Eukaryotic protein kinase domain
 NT2RP3000201//Eukaryotic protein kinase domain
 NT2RP3000232//Zinc finger, C2H2 type
 NT2RP3000304//Low-density lipoprotein receptor domain class A//Low-density lipoprotein receptor domain class
 B
 55 NT2RP3000436//Thioredoxins
 NT2RP3000460//eubacterial secY protein
 NT2RP3000616//Fibronectin type III domain
 NT2RP3000652//Zinc finger, C2H2 type

EP 1 130 094 A2

NT2RP3000676//Mitochondrial carrier proteins
 NT2RP3000789//KH domain family of RNA binding proteins
 NT2RP3000820//WD domain, G-beta repeats
 NT2RP3000838//PH (pleckstrin homology) domain
 5 NT2RP3000907//E1-E2 ATPases
 NT2RP3000921//IG superfamily
 NT2RP3001195//Sugar (and other) transporters
 NT2RP3001240//eubacterial secY protein
 NT2RP3001388//C2 domain
 10 NT2RP3001650//CUB domain//Low-density lipoprotein receptor domain class A
 NT2RP3001738//Heme-binding domain in cytochrome b5 and oxidoreductases
 NT2RP3001976//Zinc finger, C2H2 type
 NT2RP3002281//RNA recognition motif. (aka RRM, RBD, or RNP domain)
 NT2RP3002411//Alcohol/other dehydrogenases, short chain type
 15 NT2RP3002721//Citrate synthase
 NT2RP3003000//Ion transport proteins
 NT2RP3003527//Eukaryotic protein kinase domain
 NT2RP3003535//TPR Domain
 NT2RP3003849//C2 domain
 20 NT2RP3004067//Src homology domain 3
 NT2RP3004090//Zinc finger, C3HC4 type (RING finger)
 NT2RP3004481//IG superfamily
 NT2RP3004552//CUB domain//Sushi domain
 NT2RP3004647//Mitochondrial carrier proteins
 25 NT2RP4000108//Intermediate filament proteins
 NT2RP4000634//Eukaryotic protein kinase domain
 NT2RP4000962//Eukaryotic protein kinase domain
 NT2RP4001009//Zinc-binding metalloprotease domain
 NT2RP4001877//RNA recognition motif. (aka RRM, RBD, or RNP domain)
 30 NT2RP4002187//Alcohol/other dehydrogenases, short chain type
 NT2RP4002750//Amino acid permeases
 OVARC1000105//Ubiquitin-conjugating enzymes
 OVARC1000255//Eukaryotic protein kinase domain
 OVARC1000313//Thioredoxins
 35 OVARC1000410//Fibrinogen beta and gamma chains, C-terminal globular domain
 OVARC1000529//Eukaryotic protein kinase domain
 OVARC1000811//CUB domain//Kringle domain
 OVARC1000916//Eukaryotic protein kinase domain
 OVARC1001049//Helix-loop-helix DNA-binding domain
 40 OVARC1001338//Eukaryotic protein kinase domain
 OVARC1001569//Eukaryotic protein kinase domain
 OVARC1001570//Eukaryotic aspartyl proteases
 PLACE1000231//WAP-type (Whey Acidic Protein) 'four-disulfide core'
 PLACE1000258//Zinc finger, C2H2 type
 45 PLACE1000740//EGF-like domain
 PLACE1000907//Zinc finger, C2H2 type
 PLACE1001016//Ion transport proteins
 PLACE1001500//Helicases conserved C-terminal domain
 PLACE1001655//Ion transport proteins
 50 PLACE1001795//SCP-like extracellular Proteins
 PLACE1001949//E1-E2 ATPases
 PLACE1002329//Src homology domain 3
 PLACE1002355//Alpha-2-macroglobulin family//Kazal-type serine protease inhibitor domain
 PLACE1002374//Cysteine proteases
 55 PLACE1002518//Zinc finger, C3HC4 type (RING finger)
 PLACE1002911//IG superfamily
 PLACE1003135//Eukaryotic protein kinase domain
 PLACE1003163//Enoyl-CoA hydratase/isomerase

EP 1 130 094 A2

PLACE1003573//Lectin C-type domain short and long forms
 PLACE1004166//Bromodomain
 PLACE1004305//Ras family (contains ATP/GTP binding P-loop)
 PLACE1004441//7 transmembrane receptor (rhodopsin family)
 5 PLACE1004520//IG superfamily
 PLACE1004816//Fibrinogen beta and gamma chains, C-terminal globular domain
 PLACE1004887//Zinc finger, C3HC4 type (RING finger)
 PLACE1005003//Trypsin
 PLACE1005383//EGF-like domain
 10 PLACE1005410//eubacterial secY protein
 PLACE1005426//IG superfamily
 PLACE1005519//Eukaryotic protein kinase domain
 PLACE1005539//Heat shock hsp20 proteins
 PLACE1005544//IG superfamily
 15 PLACE1005569//IG superfamily
 PLACE1005682//Zinc finger, C3HC4 type (RING finger)
 PLACE1005736//PH (pleckstrin homology) domain
 PLACE1006079//Homeobox domain
 PLACE1006716//C1q domain
 20 PLACE1008282//Eukaryotic protein kinase domain
 PLACE1008549//Ets-domain
 PLACE1008744//EGF-like domain//Sushi domain
 PLACE1009067//Src homology domain 3
 PLACE1010081//Eukaryotic protein kinase domain
 25 PLACE1010251//EGF-like domain
 PLACE1010713//Alcohol/other dehydrogenases, short chain type
 PLACE1010784//7 transmembrane receptor (rhodopsin family)
 PLACE1010968//Fibronectin type III domain
 PLACE1011181//ATPases associated with various cellular activities (AAA)
 30 PLACE1011364//Eukaryotic protein kinase domain
 PLACE1011407//Zinc finger, C2H2 type
 PLACE1011708//CUB domain
 PLACE1011824//Eukaryotic protein kinase domain
 PLACE1011978//Zinc finger, C2H2 type
 35 PLACE3000181//Cadherin
 PLACE3000213//Sushi domain
 PLACB4000354//BGF-like domain//Sushi domain
 SKNMC1000082//Mitochondrial carrier proteins
 THYRO1000196//Cadherin
 40 THYRO1000580//Zinc finger, C2H2 type
 THYRO1000678//Connexin
 THYRO1000795//Mitochondrial carrier proteins
 THYRO1000956//7 transmembrane receptor (rhodopsin family)
 THYRO1001113//C2 domain
 45 THYRO1001266//Sodium:solute symporter family
 THYRO1001457//Phorbol esters / diacylglycerol binding domain//Eukaryotic protein kinase domain
 THYRO1001478//EF hand
 THYRO1001593//Eukaryotic protein kinase domain
 THYRO1001700//Eukaryotic protein kinase domain
 50 THYRO1001770//Eukaryotic protein kinase domain
 Y79AA1000030//WW/rsp5/WWP domain containing proteins
 Y79AA1000426//Transforming growth factor beta like domain
 Y79AA1000777//WD domain, G-beta repeats
 Y79AA1000876//Thioredoxins
 55 Y79AA1000967//Eukaryotic protein kinase domain
 Y79AA1001090//Ank repeat
 Y79AA1001264//DnaJ, prokaryotic heat shock protein
 Y79AA1001328//EGF-like domain

EP 1 130 094 A2

Y79AA1001427//FAD/NAD-binding domain in oxidoreductases
 Y79AA1001523//Bromodomain//Zinc finger, C3HC4 type (RING finger)
 Y79AA1001530//Tubulin
 Y79AA1001727//IG superfamily
 5 Y79AA1001787//E1-E2 ATPases
 Y79AA1001799//Mitochondrial carrier proteins
 Y79AA1002022//IG superfamily
 Y79AA1002381//Eukaryotic protein kinase domain

10 EXAMPLE 17

Functional categories based on the full-length nucleotide sequences

15 **[0243]** Prediction of functions of proteins encoded by the clones and the categorization thereof were performed based on the results of homology search (see homology search result 10) of the databases, GenBank, Swiss-Prot and UniGene for the full-length nucleotide sequences of 826 clones as well as based on the results of domain search (see Example 16) of the deduced amino acid sequences encoded by the full-length nucleotide sequences. (HEMBA1005337, NT2RM1000407, NT2RM2001767, and NT2RP3003939 were excluded because of the absence of full-length sequence.)

20 **[0244]** There are 611 clones that presumably encode proteins belonging to any of categories of secretory and/or membrane proteins, glycoprotein-associated proteins, signal transduction-associated proteins, transcription-associated proteins and disease-associated proteins.

25 **[0245]** The clones presumably encoding proteins categorized into secretory and/or membrane proteins are those which matched the full-length sequences of Swiss-Prot database with keywords "growth factor", "cytokine", "hormone", "signal", "transmembrane", "membrane", "extracellular matrix", "receptor", "G-protein coupled receptor", "ionic channel", "voltage-gated channel", "calcium channel", "cell adhesion", "collagen" or "connective tissue"; those which matched the data, suggesting that the proteins are secretory and/or membrane proteins; or those which matched the full-length sequences of GenBank or UniGene database with similar description; and, further, those predicted to have an N-terminal signal sequence or a transmembrane region as a result of domain search for the amino acid sequences deduced from the full-length nucleotide sequences.

30 **[0246]** The clones presumably encoding proteins categorized into glycoprotein-associated proteins are those which matched the full-length sequences of Swiss-Prot database with the keywords "glycoprotein"; those which matched the data, suggesting that the proteins are glycoprotein; or those which matched the full-length sequences of GenBank or UniGene database.

35 **[0247]** The clones presumably encoding proteins categorized into signal transduction-associated proteins are those which matched the full-length sequences of Swiss-Prot database with the keywords "serine/threonine-protein kinase", "tyrosine-protein kinase" or "SH3 domain"; those which matched the data, suggesting that the proteins are signal transduction-associated proteins (for example, "ADP-ribosylation factor"); or those which matched the full-length sequences of GenBank or UniGene database with similar description.

40 **[0248]** The clones presumably encoding proteins categorized into transcription-associated proteins are those which matched the full-length sequences of Swiss-Prot database with the keywords "transcription regulation", "zinc finger" or "homeobox"; those which matched the data, suggesting that the proteins are transcription-associated proteins; or those which matched the full-length sequences of GenBank or UniGene database with similar description.

45 **[0249]** The clones presumably encoding proteins categorized into disease-associated proteins are those which matched the full-length sequences of Swiss-Prot database with the keywords "disease mutation" or "syndrome"; those which matched the data, suggesting that the proteins are disease-associated proteins; or those which matched the full-length sequences of Swiss-Prot database and GenBank or UniGene database where the matched sequences are those of genes or proteins which had been deposited in the database of Online Mendelian Inheritance in Man (OMIM) (<http://www.ncbi.nlm.nih.gov/Omim/>), which is a database of human genes and diseases.

50 **[0250]** The following 437 clones were categorized into secretory and/or membrane proteins.
 BNGH41000020, BNGH41000087, BNGH41000091, HEMBA1000121, HEMBA1000128, HEMBA1000349,
 HEMBA1000477, HEMBA1000590, HEMBA1000713, HEMBA1000732, HEMBA1000745, HEMBA1000835,
 HEMBA1000940, HEMBA1000962, HEMBA1001221, HEMBA1001228, HEMBA1001621, HEMBA1002131,
 HEMBA1002163, HEMBA1002167, HEMBA1002178, HEMBA1002195, HEMBA1002227, HEMBA1002420,
 55 HEMBA1002421, HEMBA1002767, HEMBA1003047, HEMBA1003101, HEMBA1003230, HEMBA1003392,
 HEMBA1003530, HEMBA1003602, HEMBA1003732, HEMBA1003945, HEMBA1004110, HEMBA1004250,
 HEMBA1004391, HEMBA1004444, HEMBA1004454, HEMBA1004505, HEMBA1004797, HEMBA1004982,
 HEMBA1005070, HEMBA1005449, HEMBA1005522, HEMBA1005545, HEMBA1005698, HEMBA1005945,

EP 1 130 094 A2

	HEMBA1006171,	HEMBA1006299,	HEMBA1006311,	HEMBA1006335,	HEMBA1006357,	HEMBA1006430,
	HEMBA1006482,	HEMBA1006707,	HEMBA1006724,	HEMBA1006749,	HEMBA1006902,	HEMBA1006960,
	HEMBA1007241,	HEMBB1000407,	HEMBB1000447,	HEMBB1000567,	HEMBB1000679,	HEMBB1000881,
	HEMBB1001026,	HEMBB1001048,	HEMBB1001407,	HEMBB1001530,	HEMBB1001573,	HEMBB1001847,
5	HEMBB1001978,	HEMBB1002041,	HEMBB1002162,	HEMBB1002245,	HEMBB1002427,	HEMBB1002693,
	MAMMA1000102,	MAMMA1000106,	MAMMA1000118,	MAMMA1000141,	MAMMA1000204,	MAMMA1000226,
	MAMMA1000457,	MAMMA1000473,	MAMMA1000496,	MAMMA1000591,	MAMMA1000681,	MAMMA1000810,
	MAMMA1000986,	MAMMA1000994,	MAMMA1001043,	MAMMA1001141,	MAMMA1001237,	MAMMA1001344,
	MAMMA1001418,	MAMMA1001893,	MAMMA1001957,	MAMMA1001978,		
10	MAMMA1002070,	MAMMA1002091,	MAMMA1002095,	MAMMA1002165,	MAMMA1002234,	MAMMA1002586,
	MAMMA1002633,	MAMMA1003126,	NT2RM1000462,	NT2RM1000542,	NT2RM1000580,	NT2RM1000855,
	NT2RM1000858,	NT2RM1000899,	NT2RM2000241,	NT2RM2000410,	NT2RM2000423,	NT2RM2000565,
	NT2RM2001626,	NT2RM2001792,	NT2RM2001939,	NT2RM2001941,	NT2RM4000198,	NT2RM4000284,
	NT2RM4000417,	NT2RM4000444,	NT2RM4000587,	NT2RM4000593,	NT2RM4000648,	NT2RM4000761,
15	NT2RM4000997,	NT2RM4001325,	NT2RM4001735,	NT2RM4001768,	NT2RM4001843,	NT2RM4002352,
	NT2RP1000050,	NT2RP1000181,	NT2RP1000261,	NT2RP1000300,	NT2RP1000325,	NT2RP1000448,
	NT2RP1000551,	NT2RP1000613,	NT2RP1000981,	NT2RP1001004,	NT2RP1001563,	NT2RP2000479,
	NT2RP2000533,	NT2RP2000616,	NT2RP2000649,	NT2RP2000663,	NT2RP2000694,	NT2RP2000818,
	NT2RP2000903,	NT2RP2001200,	NT2RP2001276,	NT2RP2001480,	NT2RP2001495,	NT2RP2001514,
20	NT2RP2001755,	NT2RP2001915,	NT2RP2001956,	NT2RP2002063,	NT2RP2002188,	NT2RP2002232,
	NT2RP2002527,	NT2RP2002533,	NT2RP2002721,	NT2RP2002824,	NT2RP2002942,	NT2RP2002976,
	NT2RP2003042,	NT2RP2003210,	NT2RP2003383,	NT2RP2003390,	NT2RP2003469,	NT2RP2003593,
	NT2RP2003655,	NT2RP2003664,	NT2RP2003950,	NT2RP2004179,	NT2RP2004205,	NT2RP2004495,
	NT2RP2004524,	NT2RP2004556,	NT2RP2004606,	NT2RP2004648,	NT2RP2004794,	NT2RP2005027,
25	NT2RP2005163,	NT2RP2005181,	NT2RP2005378,	NT2RP2005463,	NT2RP2005541,	NT2RP2005597,
	NT2RP2005666,	NT2RP2005883,	NT2RP2005994,	NT2RP2006004,		
	NT2RP2006042,	NT2RP2006269,	NT2RP2006512,	NT2RP2006580,	NT2RP3000169,	NT2RP3000171,
	NT2RP3000304,	NT2RP3000436,	NT2RP3000460,	NT2RP3000616,	NT2RP3000676,	NT2RP3000721,
	NT2RP3000818,	NT2RP3000907,	NT2RP3000921,	NT2RP3001012,	NT2RP3001159,	NT2RP3001195,
30	NT2RP3001240,	NT2RP3001271,	NT2RP3001322,	NT2RP3001388,	NT2RP3001560,	NT2RP3001592,
	NT2RP3001650,	NT2RP3001738,	NT2RP3001858,	NT2RP3002015,	NT2RP3002160,	NT2RP3002311,
	NT2RP3002342,	NT2RP3002411,	NT2RP3002737,	NT2RP3002790,	NT2RP3002836,	NT2RP3002900,
	NT2RP3002958,	NT2RP3003000,	NT2RP3003076,	NT2RP3003354,	NT2RP3003532,	NT2RP3003535,
	NT2RP3003614,	NT2RP3004025,	NT2RP3004075,	NT2RP3004083,	NT2RP3004130,	NT2RP3004133,
35	NT2RP3004309,	NT2RP3004345,	NT2RP3004406,	NT2RP3004481,	NT2RP3004552,	NT2RP3004625,
	NT2RP3004647,	NT2RP4001001,	NT2RP4001009,	NT2RP4001467,	NT2RP4001879,	NT2RP4002187,
	NT2RP4002451,	NT2RP4002750,	OVARC1000003,	OVARC1000105,	OVARC1000298,	OVARC1000307,
	OVARC1000313,	OVARC1000410,	OVARC1000439,	OVARC1000553,	OVARC1000811,	OVARC1000873,
	OVARC1000956,	OVARC1001030,	OVARC1001163,	OVARC1001336,	OVARC1001570,	OVARC1001607,
40	OVARC1001725,	OVARC1001991,	PLACE1000033,	PLACE1000231,	PLACE1000560,	PLACE1000740,
	PLACE1000912,	PLACE1000914,	PLACE1000927,	PLACE1001016,	PLACE1001123,	PLACE1001183,
	PLACE1001231,	PLACE1001340,	PLACE1001401,	PLACE1001407,	PLACE1001464,	PLACE1001516,
	PLACE1001536,	PLACE1001564,	PLACE1001655,	PLACE1001795,		
	PLACE1001836,	PLACE1001918,	PLACE1001949,	PLACE1002080,	PLACE1002095,	PLACE1002355,
45	PLACE1002374,	PLACE1002518,	PLACE1002547,	PLACE1002726,	PLACE1002905,	PLACE1002911,
	PLACE1002967,	PLACE1003407,	PLACE1003573,	PLACE1003737,	PLACE1003772,	PLACE1003839,
	PLACE1003845,	PLACE1003852,	PLACE1004279,	PLACE1004282,	PLACE1004441,	PLACE1004450,
	PLACE1004482,	PLACE1004520,	PLACE1004630,	PLACE1004637,	PLACE1004648,	PLACE1004816,
	PLACE1005003,	PLACE1005005,	PLACE1005031,	PLACE1005383,	PLACE1005410,	PLACE1005426,
50	PLACE1005544,	PLACE1005569,	PLACE1005660,	PLACE1005725,	PLACE1005745,	PLACE1005878,
	PLACE1005927,	PLACE1006071,	PLACE1006093,	PLACE1006208,	PLACE1006277,	PLACE1006290,
	PLACE1006443,	PLACE1006716,	PLACE1006809,	PLACE1006959,	PLACE1007081,	PLACE1007096,
	PLACE1007296,	PLACE1007626,	PLACE1007845,	PLACE1007881,	PLACE1008359,	PLACE1008469,
	PLACE1008716,	PLACE1008744,	PLACE1008985,	PLACE1009067,	PLACE1009196,	PLACE1009279,
55	PLACE1009527,	PLACE1009546,	PLACE1009600,	PLACE1009982,	PLACE1010011,	PLACE1010078,
	PLACE1010251,	PLACE1010445,	PLACE1010713,	PLACE1010784,	PLACE1010827,	PLACE1010968,
	PLACE1011116,	PLACE1011181,	PLACE1011236,	PLACE1011516,	PLACE1011708,	PLACE3000181,
	PLACE3000213,	PLACE4000354,	SKNMC1000004,	SKNMC1000014,	SKNMC1000082,	THYRO1000036,

EP 1 130 094 A2

THYRO1000099, THYRO1000196, THYRO1000400, THYRO1000584, THYRO1000678, THYRO1000776,
 THYRO1000795, THYRO1000956, THYRO1001102, THYRO1001113,
 THYRO1001205, THYRO1001237, THYRO1001242, THYRO1001266, THYRO1001327, THYRO1001456,
 THYRO1001478, THYRO1001523, THYRO1001529, THYRO1001641, THYRO1001702, THYRO1001725,
 5 Y79AA1000207, Y79AA1000226, Y79AA1000270, Y79AA1000426, Y79AA1000521, Y79AA1000876,
 Y79AA1000888, Y79AA1000959, Y79AA1001013, Y79AA1001212, Y79AA1001264, Y79AA1001328,
 Y79AA1001426, Y79AA1001427, Y79AA1001430, Y79AA1001727, Y79AA1001787, Y79AA1001795,
 Y79AA1001799, Y79AA1001803, Y79AA1002022, Y79AA1002058, Y79AA1002129, Y79AA1002213,
 Y79AA1002373,
 10 **[0251]** The following 146 clones were categorized into glycoprotein-associated proteins.
 BNGH41000087, BNGH41000091, HEMBA1000349, HEMBA1000590, HEMBA1000745, HEMBA1000835,
 HEMBA1001221, HEMBA1001228, HEMBA1001621, HEMBA1002131, HEMBA1002178, HEMBA1002421,
 HEMBA1002767, HEMBA1003230, HEMBA1003392, HEMBA1004250, HEMBA1004391, HEMBA1004444,
 HEMBA1004505, HEMBA1005449, HEMBA1005522, HEMBA1005545, HEMBA1006707, HEMBA1006749,
 15 HEMBA1006902, HEMBB1000679, HEMBB1000881, HEMBB1001048, HEMBB1002120, HEMBB1002245,
 HEMBB1002427, MAMMA1000102, MAMMA1000591, MAMMA1000681, MAMMA1001043, MAMMA1001237,
 MAMMA1002070, MAMMA1002586, MAMMA1003126, NT2RM1000462, NT2RM1000580, NT2RM2001792,
 NT2RM2001818, NT2RM2001939, NT2RM2001941, NT2RM4000198, NT2RM4000284, NT2RM4000417,
 NT2RM4000648, NT2RM4000997, NT2RM4001325, NT2RM4002352, NT2RP1000613, NT2RP1000981,
 20 NT2RP1001004, NT2RP2000616, NT2RP2000694, NT2RP2000903, NT2RP2001480, NT2RP2001755,
 NT2RP2002533, NT2RP2003042, NT2RP2003210, NT2RP2004205, NT2RP2004606, NT2RP2005027,
 NT2RP2005181, NT2RP2005541, NT2RP2005597, NT2RP2005883, NT2RP2006004, NT2RP2006042,
 NT2RP2006269, NT2RP3000304, NT2RP3000616, NT2RP3000921, NT2RP3001650, NT2RP3002160,
 NT2RP3002737, NT2RP3002958, NT2RP3003000, NT2RP3003532, NT2RP3004130, NT2RP3004133,
 25 NT2RP3004481, NT2RP3004552, NT2RP3004640, NT2RP4000108, NT2RP4001467, NT2RP4002750,
 OVARC1000003, OVARC1000553, OVARC1000811, OVARC1000873, OVARC1001336, OVARC1001607,
 OVARC1001991, PLACE1000033, PLACE1000740, PLACE1001016,
 PLACE1001123, PLACE1001231, PLACE1001464, PLACE1001655, PLACE1001836, PLACE1002355,
 PLACE1002374, PLACE1002905, PLACE1002911, PLACE1003573, PLACE1003737, PLACE1003772,
 30 PLACE1003839, PLACE1004282, PLACE1004441, PLACE1004450, PLACE1004520, PLACE1004648,
 PLACE1005003, PLACE1005426, PLACE1006071, PLACE1006073, PLACE1006290, PLACE1007081,
 PLACE1007845, PLACE1008716, PLACE1008744, PLACE1008985, PLACE1010251, PLACE1010784,
 PLACE1010968, PLACE1011116, PLACE3000181, PLACE3000213, PLACE4000354, THYRO1000036,
 THYRO1000196, THYRO1000584, THYRO1000956, THYRO1001266, Y79AA1000270, Y79AA1000426,
 35 Y79AA1001727, Y79AA1001795, Y79AA1002022, Y79AA1002213,
[0252] The following 55 clones were categorized into signal transduction-associated proteins.
 HEMBA1000006, HEMBA1002195, HEMBA1002227, HEMBA1002551, HEMBA1005084, HEMBA1005929,
 HEMBA1006658, HEMBA1006916, MAMMA1000881, MAMMA1001150, MAMMA1001310, MAMMA1002142,
 NT2RM2001902, NT2RP1001020, NT2RP1001031, NT2RP2001469, NT2RP2001529, NT2RP2001769,
 40 NT2RP2003179, NT2RP2003545, NT2RP2004670, NT2RP3000011, NT2RP3000022, NT2RP3000172,
 NT2RP3000201, NT2RP3000820, NT2RP3003527, NT2RP3003849, NT2RP3003874, NT2RP3004067,
 NT2RP4000634, NT2RP4000962, OVARC1000255, OVARC1000529, OVARC1000916, OVARC1001338,
 OVARC1001569, PLACE1002329, PLACE1003135, PLACE1003598, PLACE1005519, PLACE1006208,
 PLACE1008282, PLACE1008297, PLACE1010081, PLACE1011364, PLACE1011824, THYRO1001457,
 45 THYRO1001593, THYRO1001700, THYRO1001770, Y79AA1000777, Y79AA1000967, Y79AA1002376,
 Y79AA1002381,
[0253] The following 80 clones were categorized into transcription-associated proteins.
 HEMBA1000462, HEMBA1000671, HEMBA1001297, HEMBA1001390, HEMBA1001886, HEMBA1002048,
 HEMBA1003120, HEMBA1003497, HEMBA1004785, HEMBA1005230, HEMBA1005246, HEMBA1006276,
 50 HEMBA1006572, HEMBA1007226, HEMBB1000106, HEMBB1000905, HEMBB1001959, HEMBB1002051,
 HEMBB1002661, MAMMA1001094, MAMMA1001532, MAMMA1001615, NT2RM1000789, NT2RM2000632,
 NT2RM2000773, NT2RM4000326, NT2RP1000271, NT2RP1000468, NT2RP2000092, NT2RP2000610,
 NT2RP2000712, NT2RP2000739, NT2RP2001538, NT2RP2001662, NT2RP2001817, NT2RP2001948,
 NT2RP2002564, NT2RP2002974, NT2RP2003138, NT2RP2003302, NT2RP2003940, NT2RP2004108,
 55 NT2RP2004847, NT2RP2005247, NT2RP2005391, NT2RP2005535, NT2RP2005774, NT2RP2005941,
 NT2RP2006092, NT2RP3000148, NT2RP3000232, NT2RP3000378, NT2RP3000652, NT2RP3001976,
 NT2RP3004090, NT2RP3004119, NT2RP3004294, OVARC1001049, OVARC1001086, OVARC1001132,
 OVARC1001807, PLACE1000258, PLACE1000442, PLACE1000907, PLACE1003529, PLACE1004166,

EP 1 130 094 A2

PLACE1004168, PLACE1004887, PLACE1005250, PLACE1005682, PLACE1006079, PLACE1008549,
PLACE1011407, PLACE1011978, THYRO1000580, Y79AA1000030, Y79AA1001090, Y79AA1001523,
Y79AA1002334, Y79AA1002378,

[0254] The following 85 clones were categorized into disease-associated proteins.

5 BNGH41000020, HEMBA1000349, HEMBA1000590, HEMBA1000671, HEMBA1000835, HEMBA1001184,
HEMBA1001228, HEMBA1001886, HEMBA1003120, HEMBA1004250, HEMBA1005246, HEMBA1005267,
HEMBA1006707, HEMBA1006749, HEMBA1006902, HEMBA1006916, HEMBA1007013, HEMBB1002120,
MAMMA1000204, MAMMA1002080, NT2RM2000632, NT2RM2001126, NT2RM2001558, NT2RP1000271,
NT2RP1000465, NT2RP1000579, NT2RP2000447, NT2RP2000514, NT2RP2000739, NT2RP2001223,
10 NT2RP2001529, NT2RP2001562, NT2RP2002674, NT2RP2003369, NT2RP2004108, NT2RP2004205,
NT2RP2005535, NT2RP2005941, NT2RP2006004, NT2RP3000059, NT2RP3000125, NT2RP3000201,
NT2RP3000232, NT2RP3000616, NT2RP3000677, NT2RP3000838, NT2RP3000921, NT2RP3001542,
NT2RP3002286, NT2RP3002721, NT2RP3002737, NT2RP3002738, NT2RP3004481, OVARC1000208,
OVARC1000275, OVARC1000331, OVARC1000410, OVARC1001086, OVARC1001132, OVARC1001607,
15 OVARC1001725, OVARC1001952, PLACE1000258, PLACE1000442, PLACE1000907, PLACE1001100,
PLACE1001500, PLACE1002905, PLACE1002967, PLACE1003407, PLACE1003428, PLACE1005005,
PLACE1005239, PLACE1005815, PLACE1007028, PLACE1008716, PLACE1011407, PLACE1011978,
PLACE2000118, THYRO1000580, THYRO1000866, THYRO1001071, THYRO1001478, Y79AA1001062,
Y79AA1001530,

20 **[0255]** Out of them, the following 67 clones are those which matched the data of Swiss-Prot database and GenBank or UniGene database, genes or proteins which had been deposited in the database of Online Mendelian Inheritance in Man (OMIM) (<http://www.ncbi.nlm.nih.gov/Omim/>), which is a database of human genes and diseases. (The corresponding OMIM numbers are indicated after the clone names.)

HEMBA1000349(*600046), HEMBA1000590(*603897), HEMBA1000671(*602277), HEMBA1001886(*603899),
25 HEMBA1003120(*602277), HEMBA1004250(*600976), HEMBA1005246(*602291), HEMBA1005267(*106410),
HEMBA1006707(*603897), HEMBA1006749(*603897), HEMBA1006902(*603897), HEMBA1006916(*601524),
HEMBA1007013(*603730), HEMBB1002120(*603367), MAMMA1002080(*602672), NT2RM2001126(*603785),
NT2RM2001558(*604689), NT2RP1000271(*603899), NT2RP1000465(*602231), NT2RP2000447(*602580),
NT2RP2000514(*602431), NT2RP2000739(*194558), NT2RP2001223(*603558), NT2RP2001529(*603289),
30 NT2RP2001562(*603371), NT2RP2002674(*132811), NT2RP2003369(*179555), NT2RP2004108(*601260),
NT2RP2004205(*601610), NT2RP2005535(*603899), NT2RP2006004(*600245), NT2RP3000059(*106410),
NT2RP3000125(*180202), NT2RP3000201(*604666), NT2RP3000232(*602277), NT2RP3000616(*600245),
NT2RP3000677(*142765), NT2RP3000838(*190370), NT2RP3001542(*191161), NT2RP3002286(*604331),
NT2RP3002721(*118950), NT2RP3002738(*602265), NT2RP3004481(*601610), OVARC1000208(*603603),
35 OVARC1000275(*125647), OVARC1000331(*139265), OVARC1000410(*603874), OVARC1001086(*603862),
OVARC1001725(*603046), OVARC1001952(*190370), PLACE1000258(*603971), PLACE1000442(*601260),
PLACE1000907(*194558), PLACE1001500(*603781), PLACE1002905(*125950), PLACE1003428(*603570),
PLACE1005005(*603124), PLACE1005239(*603365), PLACE1007028(*602131), PLACE1011407(*602277),
PLACE1011978(*603971), PLACE2000118(*301000), THYRO1000580(*602277), THYRO1000866(*604045),
40 THYRO1001071(*603533), Y79AA1001062(*191161), Y79AA1001530(*602662).

[0256] Out of 215 clones excluding the above-mentioned clones, HEMBB1000668 and NT2RM4001377 presumably belong to a group of signal transduction-associated proteins, based on the results of domain search by Pfam.

[0257] HEMBB1002302 presumably belong to a group of transcription-associated proteins, based on the results of domain search by Pfam.

45 **[0258]** In the 437 clones categorized into secretory and/or transmembrane proteins on the basis of their full-length sequences, 410 clones were also predicted to encode proteins having functions of secretory and/or membrane proteins on the basis of their partial nucleotide sequences (5' sequences). In the 146 clones categorized into glycoprotein-associated proteins on the basis of their full-length sequences, 124 clones were also predicted to encode proteins having functions of glycoprotein-associated proteins on the basis of their partial nucleotide sequences. In the 57 clones
50 categorized into signal transduction-associated proteins on the basis of their full-length sequences, 46 clones were also predicted to encode proteins having functions of signal transduction-associated proteins on the basis of their partial nucleotide sequences. In the 81 clones categorized into transcription-associated proteins on the basis of their full-length sequences, 57 clones were also predicted to encode proteins having functions of transcription-associated proteins on the basis of their partial nucleotide sequences. In the 85 clones categorized into disease-associated proteins
55 on the basis of their full-length sequences, 6 clones were also predicted to encode proteins having functions of disease-associated proteins on the basis of their partial nucleotide sequences. The number of clones which were predicted to encode disease-associated proteins based on the full-length nucleotide sequences is much greater than that predicted based on the partial sequences. The reason is that the full-length sequences were categorized by using the data found

in the OMIM database into the category of disease-associated proteins.

[0259] When the predicted functions based on the partial sequences were different from those based on the full-length sequences, several reasons were presumed; the ORF is too short in the partial sequence as compared with that of the full-length sequence; alternatively, P value for the partial sequence was greater than that for the full-length, that is, as compared with the probability of occurrence of the predicted function found in the full-length sequence, the probability was lower in the partial sequence. A protein does not always belong solely to a single category of the above-described functional categories, and therefore, additional functions can be found for the cDNA clones by further analyses.

[0260] It is unclear, by the analyses for the full-length sequences so far, whether or not the remaining 212 clones encode proteins belonging to any of the categories of secretory and/or membrane proteins, glycoprotein-associated proteins, signal transduction-associated proteins, transcription-associated proteins or disease-associated proteins. Nonetheless, the functions which were predicted based on the partial sequences can be verified by further analyses.

[0261] Among the 212 clones, there are 38 clones that presumably belong to the category of enzymes and/or metabolism-associated proteins, cell division- and/or cell proliferation-associated proteins, cytoskeleton-associated proteins, nuclear proteins, DNA-and/or RNA-binding proteins, ASP- and/or GTP-binding proteins, protein synthesis- and/or protein transport-associated proteins, or cellular defense-associated proteins. The clones containing results of homology search of Swiss-Prot database were categorized by considering the keywords and mentioned items in the matching data. The clones containing results of homology search of GenBank or UniGene database were categorized by considering the definitions and mentioned items in the matching data.

When the matching data contained keywords such as "metabolism", "oxidoreductase" and "E.C. No. (Enzyme commission number)", the clones were herein defined as clones presumably belonging to the category of enzymes and/or metabolism-associated proteins. When the matching data contained keywords such as "cell division", "cell cycle", "mitosis", "chromosomal protein", "cell growth" and "apoptosis", the clones were herein defined as clones presumably belonging to the category of cell division- or cell proliferation-associated proteins. When the matching data contained keywords such as "structural protein", "cytoskeleton", "actin-binding" and "microtubules", the clones were herein defined as clones presumably belonging to the category of cytoskeleton-associated proteins. When the matching data contained keywords such as "nuclear protein", the clones were herein defined as clones presumably belonging to the category of nuclear proteins. When the matching data contained keywords such as "DNA-binding" and "RNA-binding", the clones were herein defined as clones presumably belonging to the category of DNA- or RNA-binding proteins. When the matching data contained keywords such as "ATP-binding" and "GTP-binding", the clones were herein defined as clones presumably belonging to the category of ATP- and/or GTP-binding proteins. When the matching data contained keywords such as "translation regulation", "protein biosynthesis", "amino-acid biosynthesis", "ribosomal protein", "protein transport" and "signal recognition particle", the clones were herein defined as clones presumably belonging to the category of protein synthesis- and/or protein transport-associated proteins. When the matching data contained keywords such as "heat shock", "DNA repair" and "DNA damage", the clones were herein defined as clones presumably belonging to the category of cellular defense-associated proteins.

[0262] The following 10 clones presumably belong to enzymes and/or metabolism-associated proteins.

HEMBA1003315, HEMBB1002465, MAMMA1000614, NT2RP2000178, NT2RP2001388, NT2RP2001903, NT2RP2002304, NT2RP2005878, NT2RP3001685, PLACE1006219

[0263] The following 4 clones presumably belong to cell division-associated and/or cell proliferation-associated proteins.

MAMMA1000403, NT2RM2000497, NT2RP2000394, Y79AA1002121

[0264] The following 6 clones presumably belong to cytoskeleton-associated proteins.

MAMMA1001609, NT2RM2000589, NT2RP3000063, PLACE1004078, PLACE 1004492, PLACE 1008657

[0265] The following 7 clones presumably belong to nuclear proteins.

HEMBA1001878, HEMBA1002992, MAMMA1000614, NT2RM4000965, NT2RM2001738, NT2RP2001388, Y79AA1002121

[0266] The following 5 clones presumably belong to DNA- and/or RNA-binding proteins.

HEMBA1003072, HEMBA1006770, HEMBA1007332, NT2RM2000497, Y79AA1002121

[0267] The following 7 clones presumably belong to ATP- and/or GTP-binding proteins.

HEMBA1002316, MAMMA1001609, NT2RM2000306, NT2RM2000497, NT2RP2000178, NT2RP3003729, PLACE1004305

[0268] The following 7 clones presumably belong to protein synthesis- and/or protein transport-associated proteins. NT2RM4000965, NT2RP2005069, NT2RP3000481, NT2RP3000789, NT2RP4001877, OVARC1001833, OVARC1002058,

[0269] The following clone presumably belongs to cellular defense-associated proteins.

PLACE 1005539

[0270] Although it is unclear whether or not 26 out of 174 clones other than the above-mentioned clones belong to

EP 1 130 094 A2

any of the above-described categories, these clones are predicted to have some functions, based on the homology search using their full-length sequences. The clone names and the gene definitions found in the result of homology search are shown below, separated by a double-slash, //

- 5 HEMBA1000634//Homo sapiens T-cell activation protein (PGR1) gene, complete cds.
HEMBA1002524//Human MHC Class I region proline rich protein mRNA, complete cds.
HEMBA1003399//MVP1 PROTEIN.
HEMBA1005489//Mus musculus semaphorin cytoplasmic domain-associated protein 3A (Semcap3) mRNA, complete cds.
- 10 HEMBB1000542//Mus musculus bromodomain-containing protein BP75 mRNA, complete cds.
MAMMA1000788//Bos taurus P14 (p14) mRNA, complete cds.
MAMMA1002128//ABC1 PROTEIN HOMOLOG PRECURSOR.
NT2RM2000514//Homo sapiens F-box protein Fbx21 (FBX21) mRNA, complete cds.
NT2RM2000622//Mus musculus F-box protein FBL10 mRNA, partial cds.
- 15 NT2RM4000100//Homo Sapiens Leman coiled-coil protein (LCCP) mRNA, complete cds.
NT2RP2005425//Homo sapiens mRNA for AKAP450 protein.
NT2RP3001170//Mus musculus activity-dependent neuroprotective protein (Adnp) mRNA, complete cds.
NT2RP3002571//Bos taurus mRNA for lyncein.
NT2RP3004557//Human Ki nuclear autoantigen mRNA, complete cds.
- 20 OVARC1001596//Homo sapiens Arf-like 2 binding protein BART1 mRNA, complete cds.
PLACE1002153//Homo sapiens TACC2 protein (TACC2) mRNA, partial cds.
PLACE1003163//Homo sapiens DBI-related protein mRNA, complete cds.
PLACE1005736//Human mRNA for BAS-GRIP protein.
PLACE1007702//Mus musculus TRA1 mRNA, complete cds.
- 25 PLACE1011045//Homo sapiens E1-like protein mRNA, complete cds.
THYRO1000061//Mus musculus mRNA for UBE-1c1, UBE-1c2, UBE-1c3, complete cds.
THYRO1000964//Drosophila melanogaster Felle associated protein Pellino (Pli) mRNA, complete cds.
Y79AA1000776//Mus musculus mRNA for GSG1, complete cds.
Y79AA1001056//Homo sapiens MAID protein mRNA, complete cds.
- 30 Y79AA1001272//Homo sapiens retinoic acid repressible protein (RARG-1) mRNA, complete cds.
Y79AA1001793//Mus musculus mRNA for GSG1, complete cds.

[0271] So far, useful information for presuming the functions are unavailable for the remaining 148 clones, of which names are listed below.

- 35 HEMBA1000275, HEMBA1000300, HEMBA1000443, HEMBA1000875, HEMBA1000907, HEMBA1001272,
HEMBA1001296, HEMBA1001563, HEMBA1002164, HEMBA1002239, HEMBA1002985, HEMBA1003294,
HEMBA1003487, HEMBA1004007, HEMBA1004067, HEMBA1004085, HEMBA1004952, HEMBA1004971,
HEMBA1005145, HEMBA1005430, HEMBA1005913, HEMBA1006016, HEMBA1006517, HEMBA1006544,
HEMBA1006912, HEMBA1007057, HEMBA1007063, HEMBA1007291, HEMBB1000276, HEMBB1000309,
- 40 HEMBB1000642, HEMBB1001200, HEMBB1001547, HEMBB1002039, HEMBB1002228, HEMBB1002663,
MAMMA1000046, MAMMA1000449, MAMMA1000528, MAMMA1000652, MAMMA1000706, MAMMA1000814,
MAMMA1001066, MAMMA1001284, MAMMA1001623, MAMMA1001634, MAMMA1001901, MAMMA1002087,
MAMMA1002205, MAMMA1002224, NT2RM2000582, NT2RM2001643, NT2RM4000115, NT2RM4000295,
NT2RM4001321, NT2RP1000002, NT2RP1000239, NT2RP1000679, NT2RP1000740, NT2RP1000903,
- 45 NT2RP2000240, NT2RP2001878, NT2RP2001921, NT2RP2002015, NT2RP2002409, NT2RP2002510,
NT2RP2003599, NT2RP2003931, NT2RP2004069, NT2RP2004141, NT2RP2004447, NT2RP2004837,
NT2RP2005514, NT2RP2005632, NT2RP2005887, NT2RP2006099, NT2RP2006134, NT2RP3000427,
NT2RP3000444, NT2RP3000645, NT2RP3000871, NT2RP3001044, NT2RP3001061, NT2RP3001754,
NT2RP3002281, NT2RP3002324, NT2RP3002353, NT2RP3002409, NT2RP3002448, NT2RP3002664,
- 50 NT2RP3002887, NT2RP3002983, NT2RP3003448, NT2RP3003469, NT2RP3003473, NT2RP3003559,
NT2RP3003963, NT2RP3004000, NT2RP3004202, NT2RP3004321,
NT2RP3004355, NT2RP3004374, NT2RP4002715, OVARC1000090, OVARC1000137, OVARC1000467,
OVARC1000775, OVARC1000853, OVARC1000995, OVARC1001222, OVARC1001260, OVARC1001727,
OVARC1002178, PLACE1000986, PLACE1001114, PLACE1001229, PLACE1001788, PLACE1003438,
55 PLACE1003460, PLACE1003644, PLACE1004028, PLACE1004199, PLACE1004519, PLACE1005601,
PLACE1005669, PLACE1005768, PLACE1006515, PLACE1006786, PLACE1007040, PLACE1007077,
PLACE1007591, PLACE1007971, PLACE1008984, PLACE1009735, PLACE2000219, PLACE4000455,
THYRO1000846, THYRO1000999, THYRO1001063, THYRO1001128, THYRO1001471, THYRO1001495,

THYRO1001608, THYRO1001803, Y79AA1000127, Y79AA1000750, Y79AA1001592, Y79AA1001863,

EXAMPLE 18

5 Expression frequency analysis using PCR

[0272] Many genes acting at the downstream of TNF- α and IL-1, among inflammation-associated cytokines have been previously identified. The respective stimulations are transduced through independent pathways of signaling cascade. There exists another signaling cascade for both stimulations, wherein NF- κ B is a common transducing molecule shared by the two stimulations (J. Leukoc. Biol., 1994, 56(5): 542-547). It has also been revealed that many inflammation-associated genes, including IL-2, IL-6 and G-CSF, are varied in their expression levels in response to the signal through the common pathway (Trends Genet. 1999, 15(6): 229-235). A survey was performed by using ATAC-PCR method (adaptor-competitive PCR method: Nucleic Acids Res. 1997, Nov 15; 25(22): 4694-6) for genes of which expression levels were varied depending on stimulation of inflammatory cytokines, TNF- α and IL-1. It is possible that genes of which expression is varied in response to this stimulation also participate in inflammation.

[0273] Jurkat cells (Dainippon Pharmaceutical Co., Ltd.: catalog No. 06-152) were cultured in a PRMI1640 medium (Nikken Biological and Medical Institute: catalog No. 14-501F) containing 10% fetal calf serum until the cell count reached 10^7 cells. The cells were transferred into a fresh medium containing 10 ng/ml TNF- α (recombinant Tumor Necrosis Factor; Wako pure chemical Industries Inc.: catalog No. 201-13461) or IL-1 (recombinant Interleukin-1; PeprotechEC: catalog No. 200-01B) and, further, cultured at 37 $^{\circ}$ C under an atmosphere of 5% CO $_2$. The cells cultured in the presence of TNF- α were harvested 1, 3 and 7 hours after addition of TNF- α . The cells cultured in the presence of IL-1 were harvested 1 and 7 hours after addition of IL-1. Total RNA was extracted from each of the cells by AGPC method (Acid-Guanidinium-Phenol-Chloroform method: Ana Biochem. 1987, Apr; 162(1):156-9). Total RNA was also extracted from the cells in the absence of any stimulation of TNF- α and IL-1.

[0274] ATAC-PCR analysis is performed basically according to the same procedure as described in "DNA Microarray and Advanced PCR Methods" (Cell Engineering, p. 104-112, (additional volume, Genome Science Series 1), Muramatsu & Naba (eds.), Shujunnsya). Adaptor ligation reaction was performed for an internal standard sample (which was used for preparing a calibration curve for the assessment of the test samples) and test samples in the following independent two reaction systems. Combinations of each type of the 6 adaptors (AD-1, AD-2, AD-3, AD-4, AD-5, and AD-6: see the sequences shown below) with each sample are as follows:

Reaction system A

AD1: internal standard sample (x10 concentration)
AD2: sample before stimulation
AD3: internal standard sample (x3 concentration)
AD4: sample with IL-1 stimulation for 1 hour
AD5: sample with IL-1 stimulation for 7 hours
AD6: internal standard sample (x1 concentration)

Reaction system B

AD1: internal standard sample (x1 concentration)
AD2: sample with TNF stimulation for 1 hour
AD3: sample with TNF stimulation for 3 hours
AD4: internal standard sample (x3 concentration)
AD5: sample with TNF stimulation for 7 hours
AD6: internal standard sample (x10 concentration)

Adaptor sequence

AD1;

SEQ ID NO:4180//5'-GTACATATTGTCGTAGAAGCGG-3'

EP 1 130 094 A2

SEQ ID NO:4181//3'-CATGTATAACAGCAATCTTGGCCTAG-5'
AD2;
5 SEQ ID NO:4182//5'-GTACATATTGTCGTTAGAACGGGACT-3'
SEQ ID NO:4183//3'-CATGTATAACAGCAATCTTGGCCTGACTAG-5'
AD3;
10 SEQ ID NO:4184//5'-GTACATATTGTCGTTAGAACGGGCATACT-3'
SEQ ID NO:4185//3'-CATGTATAACAGCAATCTTGGCGGTATCACTAG-5'
AD4;
15 SEQ ID NO:4186//5'-GTACATATTGTCGTTAGAACGGGATCCATACT-3'
SEQ ID NO:4187//3'-CATGTATAACAGCAATCTTGGCTAGGTATGACTAG-5'
AD5;
20 SEQ ID NO:4188//5'-GTACATATTGTCGTTAGAACGGGTCAATCCATACT-3'
SEQ ID NO:4189//3'-CATGTATAACAGCAATCTTGGCAGTTAGGTATGACTAG-5'
AD6;
25 SEQ ID NO:4190//5'-GTACATATTGTCGTTAGAACGGGTACTCAATCCATACT-3'
SEQ ID NO:4191//3'-CATGTATAACAGCAATCTTGGCATGAGTTAGGTATGACTAG-5'

[0275] In this assay, the internal standard samples used were total RNA from cultured cells or human tissues from which the cDNA libraries originated. The cultured cells and the total RNAs from tissues are indicated below. Culture of the cells was performed according to the method as described in the supplier's instruction manual. RNA preparation was carried out by standard methods.

Human teratocarcinoma cell NT-2 (Stratagene, catalog No. 204101)
35 Human neuroblastoma cell SK-N-MC (Dainippon Pharmaceutical Co., Ltd., catalog No. 04-010)
Human neuroblastoma cell Y79 (Dainippon Pharmaceutical Co., Ltd., catalog No. 04-018)
Human placenta tissues total RNA (BioChin, catalog No. 064008)
Human breast tissue total RNA (Clontech, catalog No. 64037-1)

40 [0276] PCR primers used for amplification of specific genes, and names of the corresponding cDNA clones are shown below. The assay was not carried out for clones of which corresponding internal standard sample could not be prepared for the assay. The gene-specific primers were designed so that the PCR products derived from the cDNAs with adaptor were 70-200 bp in size. Sequence of the adaptor-specific primer (labeled with fluorescent dye (FAM)) used for the competitive PCR was GTACATATTGTCGTTAGAACGC (22 nucleotides, SEQ ID NO: 4192). PCR was
45 performed basically at 94. for 5 minutes; and at 94. for 30 seconds, at 50. for 60 seconds, and at 72. for 60 seconds for 30 cycles. The annealing temperature was, however, changed in some PCR experiments.

[0277] Nucleotide sequence of clone-specific primer (all the primers consist of 20 nucleotides) used in this experiment

[0278] Clone names, primer sequences, and SEQ ID NOs were shown in this order, separated with a double-slash mark, //

50

55

EP 1 130 094 A2

MAMMA1000046//GTTACATCCAAGCATACAGG//SEQ ID NO:4193
 MAMMA1000102//ACGGGGTCTCATTCTGACAC//SEQ ID NO:4194
 5 MAMMA1000141//CAAGGTAACACGAGTCTATC//SEQ ID NO:4195
 MAMMA1000226//ACTGAGGGGCAAAGGAGAGA//SEQ ID NO:4196
 MAMMA1000403//ATTTTCCTGGAGAGCCGACT//SEQ ID NO:4197
 10 MAMMA1000473//TGGAAAGTGTACCGAATTG//SEQ ID NO:4198
 MAMMA1000496//TCAATCTGGCGCTCTGTAC//SEQ ID NO:4199
 MAMMA1000614//AGTCTTTACATGCTGAGGT//SEQ ID NO:4200
 15 MAMMA1000652//TGGTGAAGACTGGGTTTGC//SEQ ID NO:4201
 MAMMA1000706//ATGGTCTTGTGGTCCAGGT//SEQ ID NO:4202
 MAMMA1000788//TGTCCAAAAGCCACACAGAG//SEQ ID NO:4203
 20 MAMMA1000810//ATACITCCCGCACCCCAAAA//SEQ ID NO:4204
 MAMMA1000814//CAGGGTTTCTGCATGTTGCC//SEQ ID NO:4205
 MAMMA1000881//ATGGAGTTTCACTCTTGTG//SEQ ID NO:4206
 25 MAMMA1000986//TGCTGCTTTACATGGGA//SEQ ID NO:4207
 MAMMA1000994//CAGGATAGAGGTTCAGGCT//SEQ ID NO:4208
 MAMMA1001066//GATGGGGTCTCACTCTGTCA//SEQ ID NO:4209
 30 MAMMA1001094//ACGTCCAGAACTACAGGCT//SEQ ID NO:4210
 MAMMA1001141//ACTGTACTTAGGATGCTTCA//SEQ ID NO:4211
 MAMMA1001237//GGGCAACCCTATGTAGATGA//SEQ ID NO:4212
 35 MAMMA1001284//GTCTGCTCTGTACATAGGG//SEQ ID NO:4213
 MAMMA1001310//ACGCCTGTAAATCCCAACCCA//SEQ ID NO:4214
 MAMMA1001344//GCCAGTTTGTCTAGGATGC//SEQ ID NO:4215
 MAMMA1001532//ACATCTATAAGGCTGTTGC//SEQ ID NO:4216
 40 MAMMA1001609//GGGTCTCACTCTGTACCCA//SEQ ID NO:4217
 MAMMA1001615//CAAGGCACACTGAGAACTGG//SEQ ID NO:4218
 MAMMA1001623//GGATTGATGCCGATCTTA//SEQ ID NO:4219
 45 MAMMA1001901//GATAGGGTCTCATTCTGTA//SEQ ID NO:4220
 MAMMA1001957//TAGTAGAGACGGGGTTTAC//SEQ ID NO:4221
 MAMMA1001978//CTCCCTCAGACGCTTTATG//SEQ ID NO:4222
 50 MAMMA1002070//GAAGAGAACTGGGGCATCC//SEQ ID NO:4223
 MAMMA1002080//ACTCTCCCTCACTACCACTG//SEQ ID NO:4224
 MAMMA1002087//AGCTCGTCTCATGGGCACT//SEQ ID NO:4225
 55 MAMMA1002091//CTGACGTAGGTGAGTCCAT//SEQ ID NO:4226

EP 1 130 094 A2

MAMMA1002095//CATGTTATGGTCAGGTAGT//SEQ ID NO:4227
 MAMMA1002128//TGCAGCAGATGAGCAAGCAG//SEQ ID NO:4228
 5 MAMMA1002142//TCTCTTAACATGCAACAGG//SEQ ID NO:4229
 MAMMA1002165//CTCAACACCCAGGCTCAAG//SEQ ID NO:4230
 MAMMA1002234//TTTGTCTTCCTCTAAACCAG//SEQ ID NO:4231
 MAMMA1002586//CTCCACCGAAAAGACCCATT//SEQ ID NO:4232
 10 MAMMA1002633//CACAGCAGCATCTCCAAGCA//SEQ ID NO:4233
 MAMMA1003126//ACAGTATCAGAGGAGCAGGA//SEQ ID NO:4234
 NT2RM1000462//AGAGCCGAGGACATTGAGG//SEQ ID NO:4235
 15 NT2RM1000542//CAAGGGCAGCTTCAGCACT//SEQ ID NO:4236
 NT2RM1000789//TTCCCTTGCTCTTCTCTGGA//SEQ ID NO:4237
 NT2RM1000855//AGCAAGGTCTCCAGAGTGTG//SEQ ID NO:4238
 NT2RM1000858//GAACAGGAAAATGCTCTCAG//SEQ ID NO:4239
 20 NT2RM2000241//CTGCTTGCCTGCCTTGTAGT//SEQ ID NO:4240
 NT2RM2000306//TAGTCCCTTCTCTGATGGTC//SEQ ID NO:4241
 NT2RM2000497//T⁻AGTAGAGACGGTGTTCA//SEQ ID NO:4242
 NT2RM2000514//TGTTCTTTTCTTTCAGCTG//SEQ ID NO:4243
 25 NT2RM2000562//GGGAATAACATCTACAACCT//SEQ ID NO:4244
 NT2RM2000588//CCCCAGAAACAGAGAAGGCT//SEQ ID NO:4245
 NT2RM2000589//TAAGGCATGTGTCTCTAAG//SEQ ID NO:4246
 30 NT2RM2000622//ATCTCGGCTATGAAGTGCTC//SEQ ID NO:4247
 NT2RM2000773//TTCTGCCCTCTCTCAACCT//SEQ ID NO:4248
 NT2RM2001126//GCAACAGCTTCTTCATGGGT//SEQ ID NO:4249
 NT2RM2001558//TCGCCCACTGCATCCTTT//SEQ ID NO:4250
 35 NT2RM2001626//TCGGGTGGCAGTGTGTGAA//SEQ ID NO:4251
 NT2RM2001643//GTAGTTCTCTCTGAAGAATC//SEQ ID NO:4252
 NT2RM2001738//CCCAGCACTTTATTTGTAG//SEQ ID NO:4253
 40 NT2RM2001792//AGTGTAGTTTGGAGATGAGA//SEQ ID NO:4254
 NT2RM2001902//TCCCCATCCAGCCACAGAAA//SEQ ID NO:4255
 NT2RM2002109//ATTGCCTATAGAAAGTCAGC//SEQ ID NO:4256
 NT2RM4000100//GCATTATAGGGCTCAAGAT//SEQ ID NO:4257
 45 NT2RM4000115//TAGTTTCGACTTC⁻GGTTCA//SEQ ID NO:4258
 NT2RM4000284//ACATCCTTCAAA⁻CAGCATC//SEQ ID NO:4259
 NT2RM4000295//GGTCTCAGGCTGCTCACAAA//SEQ ID NO:4260
 50 NT2RM4000417//⁻TGCTGTGTAGTATTCCTTAG//SEQ ID NO:4261
 NT2RM4000761//CTAATACAA⁻GCCAG⁻CAGG//SEQ ID NO:4262
 NT2RM4001377//CTAGCTTCTCTCCCAACTG//SEQ ID NO:4263
 NT2RM4001735//TCATCAGACTGCTGTAGAG//SEQ ID NO:4264
 55 NT2RM4001768//⁻TAGTAGAGGTGGGCTTCA//SEQ ID NO:4265

EP 1 130 094 A2

NT2RM4001843//CAAGGGCAGCTTTCAGCACT//SEQ ID NO:4266
 NT2RP1000239//TTAGCAGTGTATCCTTCCG//SEQ ID NO:4267
 NT2RP1000465//TTGCCCAGGCTAGTCTCGAA//SEQ ID NO:4268
 NT2RP1000468//TTGCCCAGGCTAGTCTCGAA//SEQ ID NO:4269
 NT2RP1000679//GGCCTCAGTTCCTTGCATT//SEQ ID NO:4270
 NT2RP1000740//CGCTATATTCATGGGCTT//SEQ ID NO:4271
 NT2RP1001031//CATGTACCTATACACCA//SEQ ID NO:4272
 NT2RP2000092//CACACACATGAATCATT//SEQ ID NO:4273
 NT2RP2000178//ATCTTTGTGACAGCTCCAG//SEQ ID NO:4274
 NT2RP2000240//CCATTCCACTCATACCAAGA//SEQ ID NO:4275
 NT2RP2000447//CTCTTGGCATGATGGTCTT//SEQ ID NO:4276
 NT2RP2000610//GACATGAGACAAAATTAGCC//SEQ ID NO:4277
 NT2RP2000616//AAAATAACTGCCTGGGAGGT//SEQ ID NO:4278
 NT2RP2000712//GAAGGTAGAGCTTGACAGAG//SEQ ID NO:4279
 NT2RP2000739//AGCAACAGGAATGGAGGTG//SEQ ID NO:4280
 NT2RP2000818//GGGAAGAAATGAGACAAAGA//SEQ ID NO:4281
 NT2RP2001200//ACCACATAGCTGCAGGAAAG//SEQ ID NO:4282
 NT2RP2001276//TTGCCGTTGGTGGTGGTACT//SEQ ID NO:4283
 NT2RP2001388//TGGCACAATCTCCGCTCACT//SEQ ID NO:4284
 NT2RP2001469//TAATGGGTGGTGGGAGCTGA//SEQ ID NO:4285
 NT2RP2001538//CAAAAGGCTAGGCAGCAAG//SEQ ID NO:4286
 NT2RP2001562//CACCTGCCACAAGACATT//SEQ ID NO:4287
 NT2RP2001662//AAGAGGCTGGGCAATGGCA//SEQ ID NO:4288
 NT2RP2001755//CAAGCAAATAATCCAGCCAT//SEQ ID NO:4289
 NT2RP2001817//CTAAGCAACAGAGGAATAC//SEQ ID NO:4290
 NT2RP2001921//TGTGGGTTGGTCTTTGGAA//SEQ ID NO:4291
 NT2RP2001948//GCATTGAGGACTTTCCAGA//SEQ ID NO:4292
 NT2RP2002015//GTAGTCTCTCTGAAGAATC//SEQ ID NO:4293
 NT2RP2003138//ATGGAAGAGGCTGAGGCAAA//SEQ ID NO:4294
 NT2RP2003194//CACCTCTCATGTTTCTGCAC//SEQ ID NO:4295
 NT2RP2003302//AGAGGACAGTGGGAAGATT//SEQ ID NO:4296
 NT2RP2003390//CTGTGGGTGATTTCTGGCA//SEQ ID NO:4297
 NT2RP2003593//CTGACCCCTAAGTAAATGAA//SEQ ID NO:4298
 NT2RP2003664//GGGTGTGATGTTACTTCTC//SEQ ID NO:4299
 NT2RP2003950//CTGAAGGAGGTAATGACACT//SEQ ID NO:4300
 NT2RP2004069//CATCACCAAGGTAATAGCC//SEQ ID NO:4301
 NT2RP2004108//GATAGGTTACTGGTGAATG//SEQ ID NO:4302
 NT2RP2005069//CAAAATGAACCTGTAGCCT//SEQ ID NO:4303
 NT2RP2005378//ATTCTGCCTCCCTCTTCCTC//SEQ ID NO:4304

EP 1 130 094 A2

NT2RP2005391//CAAGGTGGATTACATGGG//SEQ ID NO:4305
 NT2RP2005535//GAATTATCGCCATGTCTCT//SEQ ID NO:4306
 5 NT2RP2005597//GACCAAGACCCCTCAACCTC//SEQ ID NO:4307
 NT2RP2006092//TCCTACGTTGGTGCAGTG//SEQ ID NO:4308
 NT2RP2006134//TGACTGCCAATTAGAAACCT//SEQ ID NO:4309
 NT2RP2006208//AAAATGGCTTCTGCCTAGT//SEQ ID NO:4310
 10 NT2RP2006476//GTTACCTTCACATCCAAA//SEQ ID NO:4311
 NT2RP3000011//TGGCAAACTCAGCTCACT//SEQ ID NO:4312
 NT2RP3000031//TCTGGCTGGAGTAGTGGT//SEQ ID NO:4313
 15 NT2RP3000063//AGATAGTCACATAGACAGAG//SEQ ID NO:4314
 NT2RP3000125//ACACATCCAACCCITCACT//SEQ ID NO:4315
 NT2RP3000148//AACAGTCCAGCCCAGAAG//SEQ ID NO:4316
 NT2RP3000169//CCAAGCACGCCATATGAAGC//SEQ ID NO:4317
 20 NT2RP3000171//CAGAAATTTCCACGAGGAT//SEQ ID NO:4318
 NT2RP3000172//GGCACAGACCATCCTTGA//SEQ ID NO:4319
 NT2RP3000201//AATGGGGTTTGGCAGGTG//SEQ ID NO:4320
 25 NT2RP3000232//ACCTTCATACACTTTCCC//SEQ ID NO:4321
 NT2RP3000304//TGGTCTGCCTCATCCTTCTC//SEQ ID NO:4322
 NT2RP3000378//GAAAAGGTAGGCAGCAAG//SEQ ID NO:4323
 NT2RP3000427//ATGTAAGTGTGAGTACCC//SEQ ID NO:4324
 30 NT2RP3000444//TCTTCTCAGTCACCTCCAC//SEQ ID NO:4325
 NT2RP3000616//GTGATAGTAACACAATCCTG//SEQ ID NO:4326
 NT2RP3000645//TGGCCAGTGATGAGAGTGC//SEQ ID NO:4327
 NT2RP3000676//CAAGACAACAAACAGAAG//SEQ ID NO:4328
 35 NT2RP3000677//TCCAATTTAGCTGTACACAC//SEQ ID NO:4329
 NT2RP3000721//GAGACTGTGATGCCTTGGTG//SEQ ID NO:4330
 NT2RP3000789//AACTGATGGCTTCTGTCTCC//SEQ ID NO:4331
 40 NT2RP3000818//GGGAGAACCTAGAAAACAA//SEQ ID NO:4332
 NT2RP3000820//TCACAATCAAGCTTACGTG//SEQ ID NO:4333
 NT2RP3000838//CTAGCTTCTCTCCCACTG//SEQ ID NO:4334
 NT2RP3000871//GATTGCTGTGAGCTTGGC//SEQ ID NO:4335
 45 NT2RP3000907//GGGAGATGAAGAGGAAGCAG//SEQ ID NO:4336
 NT2RP3000921//ACATGGCTAGCACTCCTTT//SEQ ID NO:4337
 NT2RP3001012//ACCCATTCTACCTCTCTTA//SEQ ID NO:4338
 50 NT2RP3001044//AGCAAGGCAATGAGGATACT//SEQ ID NO:4339
 NT2RP3001159//AAATGTAGGGGTGGGAAGT//SEQ ID NO:4340
 NT2RP3001240//AGCAAGTCTCCAGAGTGTG//SEQ ID NO:4341
 NT2RP3001271//GTGCTTCTACCATGAGGAT//SEQ ID NO:4342
 55 NT2RP3001542//TGGTCTCACTGCCTTCAAA//SEQ ID NO:4343

EP 1 130 094 A2

NT2RP3001560//CGCCTCCACAAACAAACCCCT//SEQ ID NO:4344
 NT2RP3001592//TAAAGGGAGGATGTTCACT//SEQ ID NO:4345
 5 NT2RP3001650//GCACTTCTGGTAGTTGTCC//SEQ ID NO:4346
 NT2RP3001685//CATTCTCCTGCCTCAGCCT//SEQ ID NO:4347
 NT2RP3001754//AGTTAGTGGTGCCTGCTTCC//SEQ ID NO:4348
 NT2RP3001976//CTCACTGGCATTAGCTGGT//SEQ ID NO:4349
 10 NT2RP3002015//CAACAACCTTCCTCCTACCC//SEQ ID NO:4350
 NT2RP3002281//TAAACAGTCAACCAATGCTC//SEQ ID NO:4351
 NT2RP3002286//AAGCAAGACATTGGAGGAA//SEQ ID NO:4352
 15 NT2RP3002353//GGCTGGAACCAATTTCTGC//SEQ ID NO:4353
 NT2RP3002409//GTACCCCTAGTGAAGACCTG//SEQ ID NO:4354
 NT2RP3002411//TCAATAGCTACCTGTAGAGT//SEQ ID NO:4355
 NT2RP3002448//TCTTCTACGACATAACCA//SEQ ID NO:4356
 20 NT2RP3002571//GCCCAAAACCAACAGTAA//SEQ ID NO:4357
 NT2RP3002721//AGTGGCTTGATTTCTGTGA//SEQ ID NO:4358
 NT2RP3002737//GTGAGTACTTGCAATATCCG//SEQ ID NO:4359
 NT2RP3002738//TCCCGATGAACACCAGCTT//SEQ ID NO:4360
 25 NT2RP3002790//CTGGAGCCTGATTATGAGAA//SEQ ID NO:4361
 NT2RP3002836//CTTAGCAACATAACCTCCA//SEQ ID NO:4362
 NT2RP3002900//TTTCTCCTCCCTAACACAT//SEQ ID NO:4363
 30 NT2RP3002958//GCCCTCTTCCAGCTCTACAT//SEQ ID NO:4364
 NT2RP3002983//AATTCTCTGCTAGAAGGCT//SEQ ID NO:4365
 NT2RP3003354//GGCTGACACCTATTATCCCA//SEQ ID NO:4366
 NT2RP3003448//TGAGTCCCAGCTATTGAG//SEQ ID NO:4367
 35 NT2RP3003473//AATTATCTCTGGTAGCAC//SEQ ID NO:4368
 NT2RP3003527//TGCAATGGCCAGCTCGTAG//SEQ ID NO:4369
 NT2RP3003532//TAAACTGACTTCCTCTGGG//SEQ ID NO:4370
 40 NT2RP3003535//CTAATGCCAGTGTTCAAGA//SEQ ID NO:4371
 NT2RP3003559//AATTCTTCTCCTGCTTTGC//SEQ ID NO:4372
 NT2RP3003614//GGGCTTTGGATTGAGTGT//SEQ ID NO:4373
 NT2RP3003963//CACGATTATCTGCCAAAA//SEQ ID NO:4374
 45 NT2RP3004000//TGGCCACTGTATGAGAGGTC//SEQ ID NO:4375
 NT2RP3004025//CACATTCTGGTGGAAAAGCA//SEQ ID NO:4376
 NT2RP3004067//TGATCACACCCAGCACTTCC//SEQ ID NO:4377
 50 NT2RP3004075//TTTCAAGTCAACACCTGCAC//SEQ ID NO:4378
 NT2RP3004090//TACACTACAGATGGGCAAAA//SEQ ID NO:4379
 NT2RP3004119//TCTAAGGCTGGGTACAGTGC//SEQ ID NO:4380
 NT2RP3004130//ACATTTCTCCTCCATACGCA//SEQ ID NO:4381
 55 NT2RP3004133//TAACCCCACTATGAGGAAAG//SEQ ID NO:4382

EP 1 130 094 A2

NT2RP3004202//CAGTGGCCTGGTAATAGAGT//SEQ ID NO:4383
 NT2RP3004294//ATATTCCACTCCCCATTCCGG//SEQ ID NO:4384
 5 NT2RP3004321//GACCAACATGCTAGAAGTGC//SEQ ID NO:4385
 NT2RP3004345//AAGTCTCATTCCTAAGGTG//SEQ ID NO:4386
 NT2RP3004406//CACCTAAAAGACTAATCCCT//SEQ ID NO:4387
 NT2RP3004552//GAATCCAAAAGCCGGTAGGG//SEQ ID NO:4388
 10 NT2RP3004557//GAAAGAGGTCAAAGTACCTG//SEQ ID NO:4389
 NT2RP3004625//ATACAGGCAGCAGGAATCAC//SEQ ID NO:4390
 NT2RP3004647//CAAGACAACAAAACAGAAGG//SEQ ID NO:4391
 15 NT2RP4000634//ATGAAGAGTACCTATGTGG//SEQ ID NO:4392
 NT2RP4000962//AACCTCGGCTTGAATTCA//SEQ ID NO:4393
 NT2RP4001007//TTTCACAATGCTACAGAGG//SEQ ID NO:4394
 NT2RP4001009//AATGTCAGCGGAGCAAAAGA//SEQ ID NO:4395
 20 NT2RP4001467//GTTGAGAACTGCTGCACTGA//SEQ ID NO:4396
 NT2RP4001877//AATCATATAGTCCCAGGTG//SEQ ID NO:4397
 NT2RP4001879//GTGTGAGAGTAGTTGGGAA//SEQ ID NO:4398
 NT2RP4002187//TCAATAGCTACCTGTAGACT//SEQ ID NO:4399
 25 NT2RP4002451//CATAAACAGTGACAGCAGAA//SEQ ID NO:4400
 NT2RP4002715//ACCAAGGCAATCAGGATACT//SEQ ID NO:4401
 NT2RP4002750//CAGCATTTAGGTGTGACGAT//SEQ ID NO:4402
 30 PLACE1000040//ACATTGCTTGAGTCTTGCCA//SEQ ID NO:4403
 PLACE1000986//CAAGGAGGTAATAGGAGAT//SEQ ID NO:4404
 PLACE1002080//CTCACTCTGTCATCAGGCT//SEQ ID NO:4405
 PLACE1002547//GTACCCCTAGTGAAGACCTG//SEQ ID NO:4406
 35 PLACE1002911//CAAAGGCAAGTGTTCAGGCT//SEQ ID NO:4407
 PLACE1003407//GACTACAAAACCTGGAAAG//SEQ ID NO:4408
 PLACE1003573//GGTGGCATGTAAATAAGACT//SEQ ID NO:4409
 40 PLACE1004078//CTCAGCCTTCCAAGTAGCAG//SEQ ID NO:4410
 PLACE1004199//GGAATCTGGAGTCAAAATC//SEQ ID NO:4411
 PLACE1004305//GCACCACTTTCGTCTGAGC//SEQ ID NO:4412
 PLACE1004450//GGCACTTAGCTTCTGTGTTT//SEQ ID NO:4413
 45 PLACE1004492//TGGGCATCAATAACACCTC//SEQ ID NO:4414
 PLACE1004630//TGCTTTTGATGGCCTGGGA//SEQ ID NO:4415
 PLACE1004816//ATGTGGACAGCAAGCAGAG//SEQ ID NO:4416
 50 PLACE1005031//GGTGTTCCTTCTTGTGTTG//SEQ ID NO:4417
 PLACE1005539//ACAGGTAGAAAGAGGGAAA//SEQ ID NO:4418
 PLACE1005569//CTGAACCTGGCCTGTGAAA//SEQ ID NO:4419
 PLACE1005601//AGATGGGACTATGAACAGG//SEQ ID NO:4420
 55 PLACE1005735//CCTGTAATCCCAGCACTTG//SEQ ID NO:4421

EP 1 130 094 A2

PLACE1005815//AGAGACAGAGTTTCTCTT//SEQ ID NO:4422
 PLACE1005927//GGTGTAGCCCATGTAGAGG//SEQ ID NO:4423
 5 PLACE1006071//TTTGCTCTTTAACTCTGCCT//SEQ ID NO:4424
 PLACE1006073//TAAAGGCAGGGTGGAGGAA//SEQ ID NO:4425
 PLACE1006079//TGATGCCACTGTCTTATG//SEQ ID NO:4426
 PLACE1006786//CAAGGAGGTAATAGGAGAT//SEQ ID NO:4427
 10 PLACE1007077//ATGTGCTGGTAAGTGAATC//SEQ ID NO:4428
 PLACE1007081//GCCGCCCAAGTTACACGAA//SEQ ID NO:4429
 PLACE1007845//TAACCCCACTATGAGGAAAG//SEQ ID NO:4430
 15 PLACE1007971//GACAATCAACTGGAAGACC//SEQ ID NO:4431
 PLACE1008282//CGTGATGACGCCAACTCCA//SEQ ID NO:4432
 PLACE1008359//AACAGGGTCCCATTTATTG//SEQ ID NO:4433
 PLACE1008469//TTGTCATCCTCCTGCCTCTG//SEQ ID NO:4434
 20 PLACE1008657//CTCAGCCTTCCAAGTAGCAG//SEQ ID NO:4435
 PLACE1008744//TAAACAAGACCCAGCACCAT//SEQ ID NO:4436
 PLACE1008984//TCGAGACCGCTTCCCATAGA//SEQ ID NO:4437
 25 PLACE1009546//AGGTAGCCGATGACAAAGCC//SEQ ID NO:4438
 PLACE1010011//GGATAAAGACAAGGATGC//SEQ ID NO:4439
 PLACE1010713//TCAATAGCTACCTGTAGAGT//SEQ ID NO:4440
 PLACE1011019//AAGGAGTGAAGCTCCATGCC//SEQ ID NO:4441
 30 PLACE1011116//CCATTACAAACCCTTAACC//SEQ ID NO:4442
 PLACE1011181//TGTTATCCATTGTTACCTG//SEQ ID NO:4443
 PLACE1011364//TGTTGACGCTGATTTAGGCA//SEQ ID NO:4444
 35 PLACE3000213//TAAACAAGACCCAGCACCAT//SEQ ID NO:4445
 PLACE4000354//TAAACAAGACCCAGCACCAT//SEQ ID NO:4446
 PLACE4000455//CAAGGAGGTAATAGGAGAT//SEQ ID NO:4447
 SKNMC1000014//CGAGACAGGGTCTTGGTTTG//SEQ ID NO:4448
 40 SKNMC1000082//TTTCCTTCGGCTGGTATGGC//SEQ ID NO:4449
 Y79AA1000030//TCCTACGTTGTGGTGCAGTG//SEQ ID NO:4450
 Y79AA1000037//AGCACCCCTTAATGGAAGAT//SEQ ID NO:4451
 45 Y79AA1000127//GATGTTACTCCTCTTGGT//SEQ ID NO:4452
 Y79AA1000226//TGCTGATGCCTCTGTCCCT//SEQ ID NO:4453
 Y79AA1000270//GAAACAACCAAGCACCCCAT//SEQ ID NO:4454
 Y79AA1000750//ATAAGGGCAGCTGGGAAGTG//SEQ ID NO:4455
 50 Y79AA1000776//TGTTAGTAGCAGGAGGAAGC//SEQ ID NO:4456
 Y79AA1000777//ACAGAGTTCAGTCCCTTTA//SEQ ID NO:4457
 Y79AA1000876//GTTAGAGCACCTGGCATCA//SEQ ID NO:4458
 55 Y79AA1000888//ACTGACTGAGGAATAAGCC//SEQ ID NO:4459
 Y79AA1000959//AGAGTGAGAGCATGGGAGGT//SEQ ID NO:4460

EP 1 130 094 A2

Y79AA1000967//GGCACAGACACCATCCTTGA//SEQ ID NO:4461
Y79AA1001056//ACAAATGAGCCTGAAAAGTC//SEQ ID NO:4462
5 Y79AA1001062//TGGTCCTCACTGCCTTCAAA//SEQ ID NO:4463
Y79AA1001090//AGTGGCCTCAAAGCTCCAGT//SEQ ID NO:4464
Y79AA1001212//ACGAAAGCACTCAAATGTCA//SEQ ID NO:4465
10 Y79AA1001272//GAATGAAATGTGGTTGAGCA//SEQ ID NO:4466
Y79AA1001426//AATGATTCGGGGCAGCAGGA//SEQ ID NO:4467
Y79AA1001427//GAGACAGACACACAGAAA//SEQ ID NO:4468
15 Y79AA1001523//AGTTTTATACCAGCATTGGC//SEQ ID NO:4469
Y79AA1001530//GGTGTAGAAGTAAATGGGA//SEQ ID NO:4470
Y79AA1001592//GATTGTGTTCTTACTCCT//SEQ ID NO:4471
20 Y79AA1001727//GCTCCACCTGACGTTCTTAA//SEQ ID NO:4472
Y79AA1001795//GTCTCCCATATCGCTGTCTT//SEQ ID NO:4473
Y79AA1001803//CACTTCTAATAACCCCTTGG//SEQ ID NO:4474
25 Y79AA1001863//TTGGGATTGGAAACCGCATT//SEQ ID NO:4475
Y79AA1001874//AGAAACCACTGAGGCCAAG//SEQ ID NO:4476
Y79AA1002058//CAGAAGCAGAAGCAGGAGCA//SEQ ID NO:4477
Y79AA1002121//ATTACTGCGATTTCCTCTG//SEQ ID NO:4478
30 Y79AA1002129//GAGTTTCTTGTAGTCCA//SEQ ID NO:4479
Y79AA1002334//ATATTTGTGTTGCCTTGGG//SEQ ID NO:4480
Y79AA1002373//GGATGGCTGGGTCAAATGGT//SEQ ID NO:4481
35 Y79AA1002376//AATGATGGCTAGGGTCACTT//SEQ ID NO:4482
Y79AA1002378//TCTTCCACATTCGTACAC//SEQ ID NO:4483
Y79AA1002381//AGGGAGTAGATGTTGGTAAA//SEQ ID NO:4484
40

[0279] The result of expression frequency analysis is shown in Table 367. Only clones with correlation coefficient of 0.9 or higher are indicated in this Table. Clones that are not presented in the Table include clones for which the assay could not be performed because of low expression levels thereof in internal standard samples or because of unexpectedly smaller or larger sizes of the PCR products.

[0280] Among the clones that could be analyzed, clones of which expression levels increased by two fold in response to the IL-1_β stimulation 1 or 7 hours after the stimulation are: NT2RM2000514, NT2RP3001159, MAMMA1001237 and MAMMA1000614.

[0281] Clones of which expression levels increased by two fold in response to the TNF-stimulation 1, 3 or 7 hours after the stimulation are:

NT2RM2000582, NT2RM2002109, NT2RP1000679, NT2RP2003664, NT2RP2005597, NT2RP2004108, NT2RP3001592, NT2RP3002738, NT2RP3004133, NT2RP3004321, NT2RP3004557, NT2RP3004294, MAMMA1001237, MAMMA1000141, MAMMA1000788, MAMMA1002070, PLACE1002547, PLACE1003573, PLACE1004305, PLACE1008744, PLACE1011181, PLACE1010713, PLACE1010011, Y79AA1000776, Y79AA1002129,

[0282] Among the clones of which expression levels increased in response to IL-1_β stimulation, MAMMA1001237 was a clone of which expression level was varied in response to TNF- α stimulation. Among clones showing higher expression levels (with relative value of 5 or higher) prior to the stimulation, PLACE1002080 is an example of clones

EP 1 130 094 A2

of which expression was suppressed by the stimulation. The expression of the clone decreased by three or more fold in response to the stimulation. These genes were found to be associated with inflammatory reaction induced by IL-1, or TNF-..

[0283] In Example 15, the genes of which expression levels were varied by culturing in the presence of TNF-. were analyzed by hybridization with high-density DNA filter. As for 3 clones (NT2RP3004557, NT2RP3004294 and PLACE 1002547), the results obtained by ATAC-PCR method were similar to those obtained by hybridization method. However, the results obtained by ATAC-PCR method were not necessarily consistent with those obtained by the hybridization method. Possible reasons for the inconsistency are the difference in cells used between the two experiments, unavailability of some data in the ATAC-PCR experiment, and the difference in the method of data treatment.

Table 28

Expression of each cDNA in human tissues (The Table also contains clones with no description in Examples)

Table 185

Expression of each cDNA in human pulmonary arterial endothelial cells cultured in a medium containing bovine serum albumin, glycated bovine serum albumin or advanced glycation endproduct of bovine serum albumin (This table also contains clones without description in Examples).

In the table, EC_G_B/EC_BSA and EC_A_B/EC_BSA represent the ratios EC_glycated_BSA/EC_BSA and EC_AGE_BSA/EC_BSA, respectively.

		EC_glycated_BSA	EC_G_B	EC_A_B	
	Clone_name	EC_BSA	EC_AGE_BSA	/EC_BSA	/EC_BSA
10					
	GAPDH(Cr1)	100.81	134.21	115.16	1.33 1.14
15					
	βactin(Cr2)	1101.9	1092.57	997.36	0.99 0.91
	ADRGL1000005	26.88	38.27	36.13	1 1
	ADRGL1000007	117.89	127.25	133.21	1.08 1.13
	ADRGL1000009	29.18	25.65	26.05	1 1
20					
	ADRGL1000011	88.9	117.33	142.9	1.32 1.61
	ADRGL1000027	33.24	40.53	43.02	1.01 1.08
	ADRGL1000058	153.41	208.84	180.05	1.36 1.17
	ADRGL1000069	16.8	21.77	29.81	1 1
	ADRGL1000077	25.74	24.72	32.86	1 1
25					
	ADRGL1000092	84.52	84.15	121.76	1 1.44
	ADRGL1000099	76.19	91.53	106.01	1.2 1.39
	ADRGL1000136	52.34	44.76	63.06	0.86 1.2
	ADRGL1000147	46.08	45.18	52.15	0.98 1.13
	ADRGL1000159	31.52	40.24	42.72	1.01 1.07
30					
	ADRGL1000160	52.34	60.37	62.29	1.15 1.19
	ADRGL1000171	21.46	16.78	25.59	1 1
	ADRGL1000181	37.44	45.71	43.65	1.14 1.09
	BGGI11000015	52.42	71	65.47	1.35 1.25
35					
	BGGI11000016	127.44	122.93	147.57	0.96 1.16
	BGGI11000017	25.65	25.74	31.33	1 1
	BGGI11000022	32.82	35.19	25.56	1 1
	BGGI11000031	44.42	43.8	40.25	0.99 0.91
	BGGI11000042	120.38	146.44	165.42	1.22 1.37
40					
	BGGI11000046	74.72	58.85	84.95	0.79 1.14
	BNGH41000020	4286.08	3584.67	4330.96	0.84 1.01
	BNGH41000025	216.67	223.74	257.06	1.03 1.19
	BNGH41000026	25.76	28.16	35.52	1 1
	BNGH41000027	29.23	23.83	17.86	1 1
45					
	BNGH41000035	280.32	238.34	305.66	0.85 1.09
	BNGH41000037	59.14	54.86	54.58	0.93 0.92
	BNGH41000042	356.1	324.08	411.07	0.91 1.15
	BNGH41000048	1201.37	869.03	739.91	0.72 0.62
	BNGH41000056	33.94	31.4	40.01	1 1
50					
	BNGH41000087	77.58	81.76	91.07	1.05 1.17
	BNGH41000091	21.05	21.23	26.82	1 1
	BNGH41000157	81.11	57.28	77.46	0.71 0.95
	BNGH41000169	21.1	17.59	22.53	1 1
	BNGH41000181	63.54	56.92	70.08	0.9 1.1
55					
	BNGH41000198	32.53	26.38	34.37	1 1

EP 1 130 094 A2

	Y79AA1002373	43.96	55.06	28.34	1.25	0.91
	Y79AA1002376	3080.78	3824.05	4481.1	1.24	1.45
5	Y79AA1002378	73.33	93.61	68.22	1.28	0.93
	Y79AA1002381	248.36	288.51	304.13	1.16	1.22
	Y79AA1002388	118.82	135.82	129.37	1.14	1.09
	Y79AA1002399	36.12	30.1	32.87	1	1
10	Y79AA1002407	57.84	42.82	52.54	0.74	0.91
	Y79AA1002413	78.77	81.36	87.31	1.03	1.11
	Y79AA1002416	34.3	30.2	51.99	1	1.3
	Y79AA1002429	67.91	69.81	80.19	1.03	1.18
15	Y79AA1002431	24.66	21.16	23.98	1	1
	Y79AA1002433	27.12	18.11	23.63	1	1
	Y79AA1002445	78.66	54.58	73.75	0.69	0.94
	Y79AA1002461	29.04	24.84	32	1	1
20	Y79AA1002466	882.69	904.65	782.53	1.02	0.89
	Y79AA1002471	53.74	51.26	68.91	0.95	1.28
	Y79AA1002472	121.95	127.4	127.11	1.04	1.04
	Y79AA1002474	53.33	40.85	47.18	0.77	0.88
25	Y79AA1002482	103.36	111.11	116.07	1.07	1.12
	Y79AA1002487	30.92	25.8	32.51	1	1
	Y79AA1002490	101.4	90.92	90.54	0.9	0.89
	Y79AA1002493	107.88	125.54	105.75	1.16	0.98
30	ZRV6C1006278	46.63	30.08	32.23	0.86	0.86

Table 186

Expression of each cDNA in undifferentiated NT2 cells, in NT2 cells cultured in the presence of retinoic acid, or in NT2 cells that were cultured in the presence of retinoic acid and then further cultured in the presence of cell-division inhibitor added (This table also contains clones without description in Examples)

In the table, NT2, NT2_RA, and NT2_RA_INHIB represent untreated NT2 cells, retinoic acid-treated NT2 cells, and retinoic acid/inhibitor-treated NT2 cells, respectively. The assay was performed in triplicate (n=3), and each result was shown in the column of exp.1, exp.2, or exp.3. In addition, "t-test N/R" and "t-test N/I" represent results of test for significance of difference between the untreated cells and the retinoic acid-treated cells, and between the untreated cells and the retinoic acid/inhibitor-treated cells, respectively. The results of the test

EP 1 130 094 A2

	HEMBA1003442	4.37	4.67	4.94	3.54	6.73	5.96		
	HEMBA1003447	7.55	9.08	49.72	65.41	63.46	65.15	*	+
	HEMBA1003453	21.03	22.03	42.15	27.85	29.02	27.64		
5	HEMBA1003461	1.5	2.13	3.49	2.6	3.63	2.2		
	HEMBA1003463	2.82	3.68	6.02	5.97	3.84	6.41		
	HEMBA1003465	1.77	2.21	6.31	4.75	5.02	3.82		
	HEMBA1003480	2.58	3.91	8.62	9.63	9.6	9.42		
	HEMBA1003485	7.06	4.84	5.29	6.13	7.26	5.52		
10	HEMBA1003487	1.8	1.85	3.4	7.12	6.39	6.79	**	+
	HEMBA1003492	1.42	1.95	4.11	2.41	5.87	2.1		
	HEMBA1003494	9.36	8.61	12.16	18.24	18.69	17.83	**	+
	HEMBA1003497	2.19	2.16	3.29	3.35	6.06	2.97		
15	HEMBA1003503	0.98	1.74	3.37	5.04	3.18	2.13		
	HEMBA1003511	0.99	2.19	3.7	2.3	4.42	2.5		
	HEMBA1003528	3.33	4	6.51	5.77	5.04	4.46		
	HEMBA1003530	1.33	0.85	3.62	1.97	3.15	2.45		
	HEMBA1003531	1.14	1.72	5.39	4.74	7.24	4.51		
20	HEMBA1003532	12.97	14.66	34.3	28.69	25.31	31.26		
	HEMBA1003538	2.54	2.4	17.88	14.54	21.58	16.83		
	HEMBA1003545	0.68	2.08	3.17	1.85	3.6	2.17		
	HEMBA1003546	1.27	2.03	1.68	1.98	2.15	2.42		
	HEMBA1003548	1.4	3.18	3.6	1.41	4.15	2.23		
25	HEMBA1003553	31.29	31.45	47.99	54.36	41.34	45.65		
	HEMBA1003555	1.39	2.73	4.81	3.53	4.48	5.19		
	HEMBA1003556	1.24	1.76	2.96	3.14	5.75	3.31		
	HEMBA1003560	1.89	2.66	7.87	10.08	13.24	9.9	*	+
30	HEMBA1003565	54.27	66.88	96.28	121.29	139.88	148.68	*	+
	HEMBA1003568	1.86	2.27	3.24	2.36	7.41	2.78		
	HEMBA1003569	2.93	2.61	2.96	5.07	3.95	4.53	**	+
	HEMBA1003571	3.53	2.33	3.8	5.19	5.3	5.83	*	+
	HEMBA1003579	3.51	4.29	4.83	3.79	5.68	5.91		
35	HEMBA1003580	3.82	4.09	4.96	3.11	4.41	3.53		
	HEMBA1003581	0.82	2.62	2.07	1.63	3.19	2.4		
	HEMBA1003591	10.8	11.44	30.24	33.74	35.7	36.88	*	+
	HEMBA1003595	0.93	1.16	2.45	2.98	4.02	2.01		
	HEMBA1003597	3.15	3.18	8.74	10.82	11.39	11.59	*	+
40	HEMBA1003598	0.58	0.93	1.33	2.62	1.83	1.61	*	+
	HEMBA1003600	3.71	4.19	13.35	14.77	13.86	16.69		
	HEMBA1003602	2.84	2.64	4.89	5.89	6.97	9.14	*	+
	HEMBA1003604	2.3	3.35	5.67	6.63	8.29	8.16	*	+
	HEMBA1003610	2.33	3.2	4.48	6.12	5.64	6.81	*	+
45	HEMBA1003615	1.76	2.61	5.23	4.95	5.21	4.96		
	HEMBA1003617	3.59	3.54	8.59	6.92	11.37	8.5		
	HEMBA1003620	5.76	6.01	4.98	13.48	17.69	12.58	**	+
	HEMBA1003621	1.6	1.66	3.19	4.52	5.42	5.08	**	+
	HEMBA1003622	0.96	0.69	1.38	1.47	3.17	2.25		
50	HEMBA1003630	0.78	1.02	1.95	1.68	2.97	1.55		
	HEMBA1003637	0.66	1.93	2.59	2.11	3.11	2.63		
	HEMBA1003640	2.33	2.1	5.27	4.16	5.68	5.5		
	HEMBA1003645	1.12	1.2	4.41	2.3	3.82	3.06		
55	HEMBA1003646	0.94	1.21	1.76	1.25	3.25	1.8		
	HEMBA1003647	0.49	2.15	3.27	2.46	3.79	2.21		

Table 367

Difference in the expression level of each clone in response to TNF. stimulation or IL-1. stimulation

Before stimulation, IL-1 1h, and IL-1 7h represent relative levels of expression in the absence of the stimulation, 1 hour after the IL-1. stimulation, and 7 hours after the stimulation, respectively. TNF 1h, TNF 3h, and TNF 7h represent relative levels of expression 1 hour after the TNF. stimulation, 3 hours after the stimulation, and 7 hours after the stimulation, respectively. Correlation coefficients 1 and 2 indicate the correlation coefficients in the calibration curves prepared based on the data for the internal standard in reaction systems A and B, respectively.

Clone	IL1		TNF			Correlation coefficients			
	before stimulation	1h	7h	1h	3h	7h	1	2	
NT2RM1000858	5.6	7.6	3.8	4.7	2.1	1.7	0.98	0.94	
NT2RM1000462	0.9	0.9	0.5	0.7	0.1	0	1	1	
NT2RM1000855	1	1.3	1	1.1	0.4	0.4	1	1	
NT2RM1000789	1	0.9	0.4	1	0.4	0.6	0.96	0.98	
NT2RM2000306	0.7	1.1	0.3	1.1	0.3	0.1	1	0.98	
NT2RM2000514	0.2	0.2	0.6	0.2	0.1	0.2	0.98	0.96	
NT2RM2001126	0.5	0	0.4	0.3	0.3	1.2	0.99	0.99	
NT2RM2001902	1.3	1.6	0.6	1.3	0.8	0.8	1	1	
NT2RM2001738	1.6	1.8	1.5	1.7	0.8	0.9	0.98	1	
NT2RM2000582	0.2	0.1	0	0.7	0.1	0.1	0.99	0.99	
NT2RM2000773	1.1	1.2	1.4	2	1	0.8	0.95	1	
NT2RM2001626	0.4	0.2	0.6	0.7	0.1	0.7	1	1	
NT2RM2001643	1.6	3.1	1.2	2.4	0.7	0.8	1	1	
NT2RM2001792	0.2	0	0	0.3	0.1	0.1	0.98	0.97	
NT2RM2000589	0.2	0.1	0	0.1	0	0	1	0.99	
NT2RM2000588	0.6	0.7	0.1	0.8	0.2	0.2	1	1	
NT2RM2002109	0	0	0	0.2	0.1	0	0.99	0.99	
NT2RM4000284	6.5	9.1	4.8	10.1	3.4	3	1	1	
NT2RM4001735	3.8	4.6	2.1	5	1.6	1.4	1	1	
NT2RM4000100	0.5	0.6	0.2	0.5	0.3	0.3	0.95	0.95	
NT2RM4000417	0.2	0	0	0.2	0.1	0	0.99	0.98	
NT2RM4000761	3.2	3.2	2.2	2.6	0.7	0.7	0.95	1	
NT2RM4001843	1.5	1.8	1.7	2.8	1.2	0.6	0.98	1	
NT2RP1000239	2.1	3.2	1.2	2.1	0.5	0.6	1	0.99	
NT2RP1000465	0.9	0.3	0.3	0.9	0.2	0.1	0.97	0.96	
NT2RP1000679	0.3	0.3	0.4	0.9	0.2	0.3	0.97	1	
NT2RP1001031	1.4	1.4	0.4	1.2	0.1	0.3	1	0.98	
NT2RP2001200	2	1.5	0.8	2.2	0.7	0.7	0.99	1	
NT2RP2001562	2.7	2.4	0.7	3.6	0.4	1.1	1	0.94	
NT2RP2001948	1.1	1.5	0.7	1.3	0.6	0.7	0.97	0.99	
NT2RP2002015	1.3	1.7	0.7	1.8	0.6	0.5	0.99	1	
NT2RP2003390	2	1.7	1.3	2.3	0.6	0.5	0.99	0.99	
NT2RP2003664	0.4	0.1	0.1	0.8	0.1	0	0.99	0.99	
NT2RP2005597	1.2	1.4	0.5	2.7	2.2	2.2	0.96	0.99	
NT2RP2001469	1.7	1.4	1.2	2	0.6	0.6	1	1	
NT2RP2000240	0.9	0.9	0.3	1.4	0.7	0.3	1	1	
NT2RP2000610	2.4	2.2	2.1	2.7	1.5	1.6	0.93	0.96	

EP 1 130 094 A2

	NT2RP2001276	1	0.4	0.4	0.8	0.2	0.7	0.95	1
	NT2RP2001817	1.2	0.8	0.5	1.9	0.7	0.7	1	1
	NT2RP2004069	0.6	0.6	0.4	0.8	0.5	0.3	0.93	0.97
5	NT2RP2004108	0.3	0.2	0.6	1.1	0.4	0.5	0.96	1
	NT2RP2005391	0.7	0.5	0.1	1.2	0.3	0.3	1	0.99
	NT2RP2006092	1.6	1.2	0.9	2.1	0.6	0.7	0.97	1
	NT2RP2006134	1.2	1.5	0.7	1.9	1	0	0.91	1
	NT2RP2000818	0.9	0.3	0.3	1.6	0.3	0.3	0.95	1
10	NT2RP2000092	1.8	1.8	0.8	2	1	1	0.99	0.98
	NT2RP2000092	1.1	1.1	0.5	1.4	0.6	0.6	0.99	0.97
	NT2RP2001538	2.1	1.9	1.8	2.5	0.6	0.8	0.98	1
	NT2RP2006476	2.1	2.2	1.4	3.2	1.6	2	0.97	0.98
15	NT2RP3000616	0.1	0.1	0	0	0	0	1	1
	NT2RP3000721	2.2	2.8	0.7	2.4	0.4	0.4	1	0.98
	NT2RP3001044	1.5	1.9	0.6	2	0.7	0.4	1	1
	NT2RP3001240	0.8	1	0.8	1.5	0.6	0.7	0.97	0.99
	NT2RP3001592	0.3	0.8	0.8	1.1	0	0	0.94	0.93
20	NT2RP3002448	4.6	4.2	2.5	4.5	0.8	1.2	1	0.98
	NT2RP3002721	1.3	1.6	0.5	1.4	0.3	0.3	1	0.99
	NT2RP3002738	0.1	0	0.1	1.9	0.1	0.1	0.99	1
	NT2RP3002790	1.6	2	0.6	1.7	0.6	0.5	0.98	1
	NT2RP3002836	1.7	3	0.9	2.4	1.6	0.7	1	1
25	NT2RP3003354	0.9	0.7	0.5	0.6	0.4	0.5	0.99	0.92
	NT2RP3003614	0.5	0.4	0	0.3	0.3	0.2	0.99	0.99
	NT2RP3004075	0.8	1.4	0.7	1	0.4	0.4	1	1
	NT2RP3004130	0.3	0.4	0	0.2	0.1	0	0.93	0.96
	NT2RP3004133	1.9	3.5	0.6	3.8	1	1.3	0.99	1
30	NT2RP3004321	0.2	0.2	0	1.4	0.4	0.2	1	0.99
	NT2RP3004406	1.3	0.2	0.2	0.7	0.1	0	1	1
	NT2RP3004552	0.1	0.1	0.1	0.1	0	0	1	1
	NT2RP3004557	1.3	1.1	2.2	2.6	1.5	1.4	0.98	0.94
	NT2RP3004647	1.2	2.1	0.6	1.2	1	0.5	1	1
35	NT2RP3000201	2.3	2.9	0.4	1	1.3	0.5	1	0.98
	NT2RP3000820	1.2	1.6	0.9	1.2	0.6	0.5	1	1
	NT2RP3000818	1.4	1.5	0.7	1.8	0.5	0.7	1	0.99
	NT2RP3001159	1.2	2.5	1.2	1.4	0.6	0.7	0.99	0.99
	NT2RP3002281	1.6	2	1.2	1.8	1	1.2	0.99	1
40	NT2RP3002571	3.9	1.8	1.2	5.2	1.4	0.8	0.99	0.97
	NT2RP3002983	1.4	1.7	0.5	1.4	0.4	0.3	1	1
	NT2RP3003473	0.8	0.9	1	0.7	0.4	0.5	1	0.99
	NT2RP3001976	0.6	1.1	0.1	0.7	0.4	0.1	1	0.99
	NT2RP3002286	1.4	1.8	1	1.6	0.6	0.5	1	0.99
45	NT2RP3002353	7.7	6.4	2.2	8.7	1.1	1.3	0.94	0.99
	NT2RP3004025	1.9	2	1	2.1	1	1	0.96	0.98
	NT2RP3004119	0.8	1.1	0.4	0	0	0.2	1	0.99
	NT2RP3000171	0.7	1.3	0.6	1	0.4	0.3	0.99	1
50	NT2RP3000676	1.2	1.9	0.7	1.1	1.3	0.5	0.99	1
	NT2RP3000921	0.2	0.1	0	0.2	0.1	0	1	0.99
	NT2RP3002015	0.8	0.6	0.4	0.7	0.1	0.1	0.99	0.99
	NT2RP3004294	0	0	0	0.1	0.1	0	1	1
	NT2RP3004345	0.6	0.4	0.2	0.9	0.2	0.5	1	1
55	NT2RP3000148	1.7	2.5	0.8	2	0.8	0.8	1	1

EP 1 130 094 A2

	NT2RP3000232	0.6	0.8	0.4	0.3	0	0	1	0.99
	NT2RP3001650	2.3	1.5	1.6	1.7	1	1.3	1	1
	NT2RP3002411	0.5	0.4	0.1	0.5	0.2	0.1	1	1
5	NT2RP4001001	0.8	1.3	0.7	1	0.7	0.4	0.97	0.97
	NT2RP4001877	1.5	0.9	1.1	1.2	0.5	0.7	1	0.99
	NT2RP4002451	0.7	1	0.6	0.7	0.2	0.3	0.91	0.95
	NT2RP4000634	1	1	0.3	0.9	0.3	0.4	0.99	1
	NT2RP4002187	0.4	0.4	0.1	0.7	0.3	0.2	1	0.99
10	NT2RP4002715	1.5	1.6	0.7	1.5	0.4	0.3	1	0.99
	MAMMA1000986	3.9	4.1	1.9	4.2	1.8	1.4	0.99	1
	MAMMA1001237	0	0	1.6	0.2	0	0	0.99	0.98
	MAMMA1001978	3.5	3.4	2.3	6	3.4	2.5	0.97	0.98
	MAMMA1002080	0.4	0	0	0.4	0.1	0	1	0.99
15	MAMMA1002234	4	4.4	3	7.7	1.9	3	0.97	1
	MAMMA1000614	4.8	1	15.5	5.6	3.9	4.8	0.95	0.93
	MAMMA1000141	7.1	11.5	3.5	14.8	6.5	3.7	1	0.98
	MAMMA1000706	7.2	9.3	3.9	3.7	2.3	2.6	0.98	0.99
20	MAMMA1000788	3	3.8	2.8	8.9	4.8	4.2	0.92	0.98
	MAMMA1000994	0.3	0	0	0.4	0	0	1	1
	MAMMA1001310	4.1	6.1	3.8	8	2.5	3.6	0.99	0.95
	MAMMA1001344	2.7	4.4	2.2	3.2	2.6	2.1	1	0.99
	MAMMA1001957	2.3	2.7	1.9	1.7	1	1.8	0.99	1
25	MAMMA1002070	0.1	0.1	0	0.8	0.4	0.2	1	0.99
	MAMMA1002586	1.7	1.6	1.2	1.3	0.4	0.3	0.94	1
	MAMMA1000102	2.1	2.3	1.4	3.3	1.6	1.6	1	1
	MAMMA1001066	2.8	2.6	1.8	5.3	0.7	1.2	1	0.98
	MAMMA1001094	2.3	2.9	2	3.3	2.1	2.5	0.96	0.9
30	MAMMA1001609	2	3	1.2	2.7	1.7	2.2	0.99	0.97
	PLACE1002547	2	1.7	1.2	4.1	1.2	2	0.95	1
	PLACE1003573	0	0	0	0.1	0	0	1	0.98
	PLACE1004199	0.1	0.2	0	0	0	0	0.99	0.97
	PLACE1004305	0	0	0	0.3	0	0.2	0.96	0.99
35	PLACE1004450	0.9	0.3	0	0.1	0	0	0.98	0.98
	PLACE1005031	0.9	0	0	0.5	0	0	0.98	0.99
	PLACE1007845	0.8	1	0.4	0.4	0.1	0.1	1	0.98
	PLACE1008984	1.4	1.2	0.4	1.9	0.6	0.5	0.98	0.98
	PLACE1011116	2.6	1.5	1.6	1.6	0.3	0.4	1	1
40	PLACE1000986	0.6	0.2	0.2	0.3	0.1	0.1	1	0.98
	PLACE1004492	1.9	1.9	1.5	3.3	1	1	1	0.97
	PLACE1005569	2.6	0.4	0	1.1	0.3	0.1	0.98	0.99
	PLACE1005601	1.7	1.3	1	2.3	0.6	0.3	0.93	1
	PLACE1006079	0.6	0.3	0	0.1	0.1	0	0.98	0.99
45	PLACE1007077	1.1	0	0	0.3	0.1	0	0.97	0.98
	PLACE1008744	0.4	0.1	0.1	1.1	0.1	0	0.98	1
	PLACE1011181	0.6	0.3	0.5	1.6	0.3	0.5	0.98	0.99
	PLACE1005539	0.4	0	0.2	0.3	0.2	0	1	0.93
	PLACE1008282	1.1	0.7	0.6	1.2	0.4	0.4	0.98	1
50	PLACE1010713	0.6	0.7	0	1.4	0.5	0.4	0.99	0.95
	PLACE1010011	1.2	1.4	0.2	2.7	1.5	1.7	1	0.99
	PLACE3000213	1.9	0.2	0.1	0.8	0.1	0	0.99	1
	PLACE1002080	6.7	3.9	0.3	1.7	0.8	0.5	0.95	0.98
55	SKNMC1000082	1.3	0.1	1.1	0.7	0	0	1	1

EP 1 130 094 A2

	Y79AA1000127	1.8	1.8	1.1	2.1	0.5	0.6	1	1
	Y79AA1000226	1.4	0.8	0.6	0.9	0.3	0.4	0.99	0.99
5	Y79AA1000776	0.3	0.1	0	1.1	0.3	0.5	0.99	0.99
	Y79AA1000876	1.1	1.5	1.2	1.3	0.5	0.8	0.97	1
	Y79AA1001056	1.7	1.7	0.8	1.4	0.9	0.7	1	1
	Y79AA1000777	3.1	3.1	1.2	3.8	0.7	0.5	0.98	0.99
	Y79AA1000030	1	1.3	0.2	1.3	0	0.6	0.98	0.96
10	Y79AA1001212	1.5	1.2	1	2	0.8	0.5	1	0.99
	Y79AA1001427	2.3	3	0.6	2	0.8	0.4	1	1
	Y79AA1001530	0.9	0.9	0.5	1.1	0.4	0.4	1	1
	Y79AA1001592	0.6	0.2	0	0.7	0	0	0.97	1
15	Y79AA1001727	0.8	0.4	0.2	0.9	0.2	0.1	1	1
	Y79AA1001803	0.1	0	0	0.2	0.1	0	0.97	0.99
	Y79AA1002373	0	0	0	0	0	0	0.99	1
	Y79AA1002376	0.9	0.1	0	1.2	0.1	0.4	0.98	1
20	Y79AA1001523	0.5	0.5	0.3	0.6	0.3	0.1	1	0.98
	Y79AA1000888	1.1	1	0.7	1.4	0.7	0.5	1	1
	Y79AA1002129	0.2	0.2	0.1	0.5	0.2	0.2	0.99	1

25 **[0285]** The present invention has provided a total of 830 novel full length cDNA clones. As has not yet proceeded the isolation of full length cDNA within the human, the invention has a large significance. Those proteins such as secretory proteins, membrane proteins, and proteins associated with signal transduction, glycoprotein, and transcription are known to be associated with many diseases. Those genes and proteins associating with diseases are useful for developing medicines as they can be used as a diagnostic marker, or a target for gene therapy or developing

30 medicines that is capable of regulating their expression and activity. Especially, the cDNA clones encoding a secretion protein are extremely important for medicinal industry since the protein itself is expected to be effective as a medicine, and also the gene may have potential to be associating with many diseases. Moreover, those proteins such as membrane proteins, and proteins associated with signal transduction, glycoprotein, transcription, and diseases, and the genes encoding the proteins may be used as a disease marker. These cDNA clones are also important for medicinal

35 industry as they may be effective for treating diseases through the regulation of the expression and activity of their encoded proteins.

Table 368

40 The names of the representative sequences of the clusters (groups) and the corresponding SEQ IDs.

	HRIFA000016a : 1573	HRIFA017855a : 1979
	HRIFA000071a : 1574	HRIFA017921a : 1980
45	HRIFA000116a : 1575	HRIFA013075a : 1981
	HRIFA000123a : 1576	HRIFA013092a : 1982
	HRIFA000264a : 1577	HRIFA013131a : 1983
	HRIFA000284a : 1578	HRIFA013134a : 1984
50	HRIFA000327a : 1579	HRIFA018238a : 1985
	HRIFA000415a : 1580	HRIFA018262a : 1986
	HRIFA000432a : 1581	HRIFA018287a : 1987
	HRIFA000446a : 1582	HRIFA018447a : 1988
55	HRIFA000553a : 1583	HRIFA018580a : 1989
	HRIFA000564a : 1584	HRIFA018666a : 1990
	HRIFA000631a : 1585	HRIFA018688a : 1991

EP 1 130 094 A2

	HRIFA000683a : 1586	HRIFA018754a : 1992
	HRIFA000695a : 1587	HRIFA018794a : 1993
	HRIFA000776a : 1588	HRIFA018827a : 1994
5	HRIFA000814a : 1589	HRIFA018849a : 1995
	HRIFA000822a : 1590	HRIFA018870a : 1996
	HRIFA000845a : 1591	HRIFA018904a : 1997
	HRIFA000899a : 1592	HRIFA018931a : 1998
	HRIFA000974a : 1593	HRIFA018993a : 1999
10	HRIFA001099a : 1594	HRIFA019105a : 2000
	HRIFA001132a : 1595	HRIFA019136a : 2001
	HRIFA001138a : 1596	HRIFA019175a : 2002
	HRIFA001179a : 1597	HRIFA019185a : 2003
15	HRIFA001200a : 1598	HRIFA019262a : 2004
	HRIFA001201a : 1599	HRIFA019412a : 2005
	HRIFA001337a : 1600	HRIFA019437a : 2006
	HRIFA001341a : 1601	HRIFA019466a : 2007
	HRIFA001413a : 1602	HRIFA019490a : 2008
20	HRIFA001439a : 1603	HRIFA019498a : 2009
	HRIFA001489a : 1604	HRIFA019532a : 2010
	HRIFA001558a : 1605	HRIFA019551a : 2011
	HRIFA001712a : 1606	HRIFA019867a : 2012
	HRIFA001720a : 1607	HRIFA019869a : 2013
25	HRIFA001866a : 1608	HRIFA019958a : 2014
	HRIFA001942a : 1609	HRIFA019960a : 2015
	HRIFA001971a : 1610	HRIFA020109a : 2016
	HRIFA001972a : 1611	HRIFA020144a : 2017
	HRIFA001975a : 1612	HRIFA020163a : 2018
30	HRIFA001984a : 1613	HRIFA020184a : 2019
	HRIFA002037a : 1614	HRIFA020272a : 2020
	HRIFA002063a : 1615	HRIFA020335a : 2021
	HRIFA002102a : 1616	HRIFA020349a : 2022
	HRIFA002195a : 1617	HRIFA020453a : 2023
35	HRIFA002284a : 1618	HRIFA020466a : 2024
	HRIFA002309a : 1619	HRIFA020661a : 2025
	HRIFA002384a : 1620	HRIFA020693a : 2026
	HRIFA002503a : 1621	HRIFA020707a : 2027
	HRIFA002689a : 1622	HRIFA020748a : 2028
40	HRIFA002694a : 1623	HRIFA020862a : 2029
	HRIFA002743a : 1624	HRIFA020883a : 2030
	HRIFA002762a : 1625	HRIFA020965a : 2031
	HRIFA002766a : 1626	HRIFA021007a : 2032
	HRIFA002787a : 1627	HRIFA021040a : 2033
45	HRIFA002805a : 1628	HRIFA021061a : 2034
	HRIFA002891a : 1629	HRIFA021069a : 2035
	HRIFA002919a : 1630	HRIFA021136a : 2036
	HRIFA002980a : 1631	HRIFA021213a : 2037
50	HRIFA002985a : 1632	HRIFA021224a : 2038
	HRIFA003055a : 1633	HRIFA021323a : 2039
	HRIFA003063a : 1634	HRIFA021398a : 2040
	HRIFA003093a : 1635	HRIFA021399a : 2041
	HRIFA003169a : 1636	HRIFA021445a : 2042
55	HRIFA003340a : 1637	HRIFA021448a : 2043

EP 1 130 094 A2

	HRIFA003357a : 1638	HRIFA021494a : 2044
	HRIFA003379a : 1639	HRIFA021499a : 2045
	HRIFA003402a : 1640	HRIFA021543a : 2046
5	HRIFA003504a : 1641	HRIFA021594a : 2047
	HRIFA003592a : 1642	HRIFA021611a : 2048
	HRIFA003635a : 1643	HRIFA021620a : 2049
	HRIFA003640a : 1644	HRIFA021637a : 2050
10	HRIFA003883a : 1645	HRIFA021651a : 2051
	HRIFA003892a : 1646	HRIFA021754a : 2052
	HRIFA003946a : 1647	HRIFA021781a : 2053
	HRIFA004006a : 1648	HRIFA021787a : 2054
	HRIFA004034a : 1649	HRIFA021794a : 2055
15	HRIFA004112a : 1650	HRIFA021855a : 2056
	HRIFA004145a : 1651	HRIFA021886a : 2057
	HRIFA004162a : 1652	HRIFA021906a : 2058
	HRIFA004401a : 1653	HRIFA021985a : 2059
	HRIFA004426a : 1654	HRIFA022055a : 2060
20	HRIFA004490a : 1655	HRIFA022065a : 2061
	HRIFA004523a : 1656	HRIFA022106a : 2062
	HRIFA004663a : 1657	HRIFA022139a : 2063
	HRIFA004696a : 1658	HRIFA022156a : 2064
	HRIFA004714a : 1659	HRIFA022166a : 2065
25	HRIFA004745a : 1660	HRIFA022177a : 2066
	HRIFA004780a : 1661	HRIFA022182a : 2067
	HRIFA004852a : 1662	HRIFA022203a : 2068
	HRIFA004919a : 1663	HRIFA022227a : 2069
30	HRIFA005072a : 1664	HRIFA022234a : 2070
	HRIFA005102a : 1665	HRIFA022249a : 2071
	HRIFA005184a : 1666	HRIFA022262a : 2072
	HRIFA005214a : 1667	HRIFA022265a : 2073
	HRIFA005231a : 1668	HRIFA022328a : 2074
35	HRIFA005240a : 1669	HRIFA022335a : 2075
	HRIFA005255a : 1670	HRIFA022348a : 2076
	HRIFA005271a : 1671	HRIFA022411a : 2077
	HRIFA005296a : 1672	HRIFA022423a : 2078
	HRIFA005300a : 1673	HRIFA022462a : 2079
40	HRIFA005356a : 1674	HRIFA022493a : 2080
	HRIFA005369a : 1675	HRIFA022528a : 2081
	HRIFA005372a : 1676	HRIFA022546a : 2082
	HRIFA005392a : 1677	HRIFA022564a : 2083
	HRIFA005409a : 1678	HRIFA022616a : 2084
45	HRIFA005420a : 1679	HRIFA022671a : 2085
	HRIFA005438a : 1680	HRIFA022691a : 2086
	HRIFA005462a : 1681	HRIFA022702a : 2087
	HRIFA005500a : 1682	HRIFA022707a : 2088
50	HRIFA005540a : 1683	HRIFA022714a : 2089
	HRIFA005644a : 1684	HRIFA022728a : 2090
	HRIFA005650a : 1685	HRIFA022729a : 2091
	HRIFA005702a : 1686	HRIFA022737a : 2092
	HRIFA005720a : 1687	HRIFA022776a : 2093
55	HRIFA005728a : 1688	HRIFA022782a : 2094
	HRIFA005732a : 1689	HRIFA022794a : 2095

EP 1 130 094 A2

	HRIFA005760a : 1690	HRIFA022865a : 2096
	HRIFA005781a : 1691	HRIFA022875a : 2097
	HRIFA005944a : 1692	HRIFA022890a : 2098
5	HRIFA006018a : 1693	HRIFA022895a : 2099
	HRIFA006183a : 1694	HRIFA022985a : 2100
	HRIFA006216a : 1695	HRIFA023007a : 2101
	HRIFA006250a : 1696	HRIFA023048a : 2102
	HRIFA006298a : 1697	HRIFA023069a : 2103
10	HRIFA006448a : 1698	HRIFA023129a : 2104
	HRIFA006494a : 1699	HRIFA023154a : 2105
	HRIFA006510a : 1700	HRIFA023212a : 2106
	HRIFA006566a : 1701	HRIFA023227a : 2107
15	HRIFA006572a : 1702	HRIFA023257a : 2108
	HRIFA006586a : 1703	HRIFA023304a : 2109
	HRIFA006596a : 1704	HRIFA023305a : 2110
	HRIFA006609a : 1705	HRIFA023434a : 2111
	HRIFA006633a : 1706	HRIFA023464a : 2112
20	HRIFA006642a : 1707	HRIFA023489a : 2113
	HRIFA006649a : 1708	HRIFA023521a : 2114
	HRIFA006667a : 1709	HRIFA023605a : 2115
	HRIFA006730a : 1710	HRIFA023619a : 2116
	HRIFA006798a : 1711	HRIFA023634a : 2117
25	HRIFA006926a : 1712	HRIFA023767a : 2118
	HRIFA007013a : 1713	HRIFA023894a : 2119
	HRIFA007026a : 1714	HRIFA023923a : 2120
	HRIFA007032a : 1715	HRIFA024088a : 2121
30	HRIFA007068a : 1716	HRIFA024132a : 2122
	HRIFA007152a : 1717	HRIFA024185a : 2123
	HRIFA007219a : 1718	HRIFA024197a : 2124
	HRIFA007228a : 1719	HRIFA024218a : 2125
	HRIFA007243a : 1720	HRIFA024255a : 2126
35	HRIFA037838a : 1721	HRIFA024305a : 2127
	HRIFA007256a : 1722	HRIFA024392a : 2128
	HRIFA007262a : 1723	HRIFA024423a : 2129
	HRIFA007352a : 1724	HRIFA024473a : 2130
	HRIFA007424a : 1725	HRIFA024482a : 2131
40	HRIFA007435a : 1726	HRIFA024504a : 2132
	HRIFA007463a : 1727	HRIFA024533a : 2133
	HRIFA007493a : 1728	HRIFA024543a : 2134
	HRIFA007512a : 1729	HRIFA024554a : 2135
45	HRIFA007532a : 1730	HRIFA024718a : 2136
	HRIFA007541a : 1731	HRIFA024767a : 2137
	HRIFA007547a : 1732	HRIFA024841a : 2138
	HRIFA007565a : 1733	HRIFA024884a : 2139
	HRIFA007571a : 1734	HRIFA024893a : 2140
50	HRIFA007619a : 1735	HRIFA024921a : 2141
	HRIFA007659a : 1736	HRIFA024937a : 2142
	HRIFA007722a : 1737	HRIFA024978a : 2143
	HRIFA007728a : 1738	HRIFA024994a : 2144
	HRIFA007745a : 1739	HRIFA025033a : 2145
55	HRIFA007829a : 1740	HRIFA025046a : 2146
	HRIFA007909a : 1741	HRIFA025125a : 2147

EP 1 130 094 A2

	HRIFA007985a : 1742	HRIFA025143a : 2148
	HRIFA008000a : 1743	HRIFA025250a : 2149
	HRIFA008099a : 1744	HRIFA025261a : 2150
5	HRIFA008131a : 1745	HRIFA025276a : 2151
	HRIFA034865a : 1746	HRIFA025290a : 2152
	HRIFA008186a : 1747	HRIFA025327a : 2153
	HRIFA008200a : 1748	HRIFA025353a : 2154
	HRIFA008212a : 1749	HRIFA025360a : 2155
10	HRIFA008252a : 1750	HRIFA025479a : 2156
	HRIFA008284a : 1751	HRIFA025488a : 2157
	HRIFA008314a : 1752	HRIFA025492a : 2158
	HRIFA008362a : 1753	HRIFA025498a : 2159
15	HRIFA008387a : 1754	HRIFA025514a : 2160
	HRIFA008426a : 1755	HRIFA025636a : 2161
	HRIFA008459a : 1756	HRIFA025695a : 2162
	HRIFA008483a : 1757	HRIFA025703a : 2163
	HRIFA008547a : 1758	HRIFA025706a : 2164
20	HRIFA008596a : 1759	HRIFA025766a : 2165
	HRIFA008606a : 1760	HRIFA025771a : 2166
	HRIFA008611a : 1761	HRIFA025778a : 2167
	HRIFA008661a : 1762	HRIFA025786a : 2168
	HRIFA008717a : 1763	HRIFA025792a : 2169
25	HRIFA008753a : 1764	HRIFA025300a : 2170
	HRIFA008784a : 1765	HRIFA025904a : 2171
	HRIFA008790a : 1766	HRIFA025907a : 2172
	HRIFA008827a : 1767	HRIFA037340a : 2173
	HRIFA008949a : 1768	HRIFA025936a : 2174
30	HRIFA008976a : 1769	HRIFA025966a : 2175
	HRIFA008981a : 1770	HRIFA025978a : 2176
	HRIFA008989a : 1771	HRIFA026089a : 2177
	HRIFA009071a : 1772	HRIFA026121a : 2178
35	HRIFA009101a : 1773	HRIFA026151a : 2179
	HRIFA009123a : 1774	HRIFA026153a : 2180
	HRIFA009136a : 1775	HRIFA026203a : 2181
	HRIFA009171a : 1776	HRIFA026210a : 2182
	HRIFA009214a : 1777	HRIFA026242a : 2183
40	HRIFA009220a : 1778	HRIFA026265a : 2184
	HRIFA009339a : 1779	HRIFA026303a : 2185
	HRIFA009372a : 1780	HRIFA026316a : 2186
	HRIFA009392a : 1781	HRIFA026346a : 2187
	HRIFA009414a : 1782	HRIFA026351a : 2188
45	HRIFA009451a : 1783	HRIFA026364a : 2189
	HRIFA009482a : 1784	HRIFA026382a : 2190
	HRIFA009578a : 1785	HRIFA026465a : 2191
	HRIFA009762a : 1786	HRIFA026490a : 2192
	HRIFA009764a : 1787	HRIFA026496a : 2193
50	HRIFA009783a : 1788	HRIFA026519a : 2194
	HRIFA009825a : 1789	HRIFA026564a : 2195
	HRIFA009852a : 1790	HRIFA026576a : 2196
	HRIFA009881a : 1791	HRIFA026615a : 2197
	HRIFA009983a : 1792	HRIFA026618a : 2198
55	HRIFA010005a : 1793	HRIFA026659a : 2199

EP 1 130 094 A2

	HRIFA010034a : 1794	HRIFA026735a : 2200
	HRIFA010070a : 1795	HRIFA026764a : 2201
	HRIFA010078a : 1796	HRIFA026789a : 2202
5	HRIFA010085a : 1797	HRIFA026813a : 2203
	HRIFA010090a : 1798	HRIFA026860a : 2204
	HRIFA010130a : 1799	HRIFA026889a : 2205
	HRIFA010152a : 1800	HRIFA026899a : 2206
	HRIFA010176a : 1801	HRIFA026918a : 2207
10	HRIFA010206a : 1802	HRIFA026923a : 2208
	HRIFA010301a : 1803	HRIFA027012a : 2209
	HRIFA010319a : 1804	HRIFA027045a : 2210
	HRIFA010322a : 1805	HRIFA027125a : 2211
15	HRIFA010361a : 1806	HRIFA027173a : 2212
	HRIFA010394a : 1807	HRIFA027179a : 2213
	HRIFA010425a : 1808	HRIFA027187a : 2214
	HRIFA010460a : 1809	HRIFA027265a : 2215
	HRIFA010466a : 1810	HRIFA027327a : 2216
20	HRIFA010490a : 1811	HRIFA027329a : 2217
	HRIFA010593a : 1812	HRIFA027355a : 2218
	HRIFA010736a : 1813	HRIFA027485a : 2219
	HRIFA010790a : 1814	HRIFA027493a : 2220
	HRIFA010799a : 1815	HRIFA027536a : 2221
25	HRIFA010859a : 1816	HRIFA027549a : 2222
	HRIFA010891a : 1817	HRIFA027619a : 2223
	HRIFA010942a : 1818	HRIFA027622a : 2224
	HRIFA010975a : 1819	HRIFA027625a : 2225
	HRIFA010988a : 1820	HRIFA027644a : 2226
30	HRIFA011016a : 1821	HRIFA027656a : 2227
	HRIFA011105a : 1822	HRIFA027673a : 2228
	HRIFA011128a : 1823	HRIFA027681a : 2229
	HRIFA011179a : 1824	HRIFA027701a : 2230
35	HRIFA011193a : 1825	HRIFA027714a : 2231
	HRIFA011197a : 1826	HRIFA027717a : 2232
	HRIFA011347a : 1827	HRIFA027722a : 2233
	HRIFA011403a : 1828	HRIFA027754a : 2234
	HRIFA011449a : 1829	HRIFA027803a : 2235
40	HRIFA011484a : 1830	HRIFA027860a : 2236
	HRIFA011512a : 1831	HRIFA027867a : 2237
	HRIFA011580a : 1832	HRIFA027874a : 2238
	HRIFA011659a : 1833	HRIFA027940a : 2239
	HRIFA011802a : 1834	HRIFA027961a : 2240
45	HRIFA011820a : 1835	HRIFA028061a : 2241
	HRIFA011926a : 1836	HRIFA028157a : 2242
	HRIFA011947a : 1837	HRIFA028187a : 2243
	HRIFA012022a : 1838	HRIFA028262a : 2244
	HRIFA012069a : 1839	HRIFA028371a : 2245
50	HRIFA012151a : 1840	HRIFA028401a : 2246
	HRIFA012167a : 1841	HRIFA028402a : 2247
	HRIFA012278a : 1842	HRIFA028440a : 2248
	HRIFA012282a : 1843	HRIFA028465a : 2249
	HRIFA012283a : 1844	HRIFA028468a : 2250
55	HRIFA012290a : 1845	HRIFA028497a : 2251

EP 1 130 094 A2

	HRIFA012333a : 1846	HRIFA028501a : 2252
	HRIFA012354a : 1847	HRIFA028511a : 2253
	HRIFA012417a : 1848	HRIFA028573a : 2254
5	HRIFA012427a : 1849	HRIFA028576a : 2255
	HRIFA012436a : 1850	HRIFA028592a : 2256
	HRIFA012480a : 1851	HRIFA028614a : 2257
	HRIFA012513a : 1852	HRIFA028651a : 2258
	HRIFA012515a : 1853	HRIFA028652a : 2259
10	HRIFA012584a : 1854	HRIFA028654a : 2260
	HRIFA012625a : 1855	HRIFA028708a : 2261
	HRIFA012692a : 1856	HRIFA028790a : 2262
	HRIFA012702a : 1857	HRIFA028804a : 2263
	HRIFA012737a : 1858	HRIFA028867a : 2264
15	HRIFA012761a : 1859	HRIFA028911a : 2265
	HRIFA012795a : 1860	HRIFA028926a : 2266
	HRIFA012881a : 1861	HRIFA028983a : 2267
	HRIFA012885a : 1862	HRIFA029002a : 2268
20	HRIFA012914a : 1863	HRIFA029050a : 2269
	HRIFA012969a : 1864	HRIFA029107a : 2270
	HRIFA012990a : 1865	HRIFA029208a : 2271
	HRIFA012999a : 1866	HRIFA029209a : 2272
	HRIFA013092a : 1867	HRIFA029256a : 2273
25	HRIFA013103a : 1868	HRIFA029263a : 2274
	HRIFA013135a : 1869	HRIFA029274a : 2275
	HRIFA013235a : 1870	HRIFA029278a : 2276
	HRIFA013254a : 1871	HRIFA029285a : 2277
	HRIFA013265a : 1872	HRIFA029317a : 2278
30	HRIFA013276a : 1873	HRIFA029327a : 2279
	HRIFA013279a : 1874	HRIFA029349a : 2280
	HRIFA013288a : 1875	HRIFA029393a : 2281
	HRIFA013376a : 1876	HRIFA029398a : 2282
35	HRIFA013477a : 1877	HRIFA029425a : 2283
	HRIFA013586a : 1878	HRIFA029434a : 2284
	HRIFA013589a : 1879	HRIFA029440a : 2285
	HRIFA013620a : 1880	HRIFA029460a : 2286
	HRIFA013668a : 1881	HRIFA029467a : 2287
40	HRIFA013726a : 1882	HRIFA029508a : 2288
	HRIFA013744a : 1883	HRIFA029511a : 2289
	HRIFA013899a : 1884	HRIFA029577a : 2290
	HRIFA013911a : 1885	HRIFA029602a : 2291
	HRIFA013919a : 1886	HRIFA029649a : 2292
45	HRIFA013932a : 1887	HRIFA029715a : 2293
	HRIFA013980a : 1888	HRIFA029730a : 2294
	HRIFA014006a : 1889	HRIFA029779a : 2295
	HRIFA014024a : 1890	HRIFA029792a : 2296
	HRIFA014056a : 1891	HRIFA029802a : 2297
50	HRIFA014111a : 1892	HRIFA029866a : 2298
	HRIFA014133a : 1893	HRIFA029932a : 2299
	HRIFA014178a : 1894	HRIFA030025a : 2300
	HRIFA014185a : 1895	HRIFA030045a : 2301
	HRIFA014303a : 1896	HRIFA030097a : 2302
55	HRIFA014336a : 1897	HRIFA030103a : 2303

EP 1 130 094 A2

	HRIFA014396a : 1898	HRIFA030106a : 2304
	HRIFA014397a : 1899	HRIFA030147a : 2305
	HRIFA014417a : 1900	HRIFA030203a : 2306
5	HRIFA014465a : 1901	HRIFA030237a : 2307
	HRIFA014467a : 1902	HRIFA030248a : 2308
	HRIFA014482a : 1903	HRIFA030250a : 2309
	HRIFA014500a : 1904	HRIFA030264a : 2310
	HRIFA014561a : 1905	HRIFA030342a : 2311
10	HRIFA014568a : 1906	HRIFA030370a : 2312
	HRIFA014590a : 1907	HRIFA030371a : 2313
	HRIFA014598a : 1908	HRIFA030381a : 2314
	HRIFA014620a : 1909	HRIFA030385a : 2315
	HRIFA014621a : 1910	HRIFA030394a : 2316
15	HRIFA014688a : 1911	HRIFA030408a : 2317
	HRIFA014692a : 1912	HRIFA030411a : 2318
	HRIFA014702a : 1913	HRIFA030448a : 2319
	HRIFA014819a : 1914	HRIFA030456a : 2320
20	HRIFA014868a : 1915	HRIFA030461a : 2321
	HRIFA014951a : 1916	HRIFA030472a : 2322
	HRIFA014953a : 1917	HRIFA030509a : 2323
	HRIFA014967a : 1918	HRIFA030511a : 2324
	HRIFA015063a : 1919	HRIFA030545a : 2325
25	HRIFA015070a : 1920	HRIFA030566a : 2326
	HRIFA015122a : 1921	HRIFA030599a : 2327
	HRIFA015129a : 1922	HRIFA030629a : 2328
	HRIFA015219a : 1923	HRIFA030642a : 2329
	HRIFA015236a : 1924	HRIFA030662a : 2330
30	HRIFA015246a : 1925	HRIFA030839a : 2331
	HRIFA015351a : 1926	HRIFA030981a : 2332
	HRIFA015409a : 1927	HRIFA031062a : 2333
	HRIFA015423a : 1928	HRIFA031075a : 2334
	HRIFA015443a : 1929	HRIFA031091a : 2335
35	HRIFA015453a : 1930	HRIFA031126a : 2336
	HRIFA015471a : 1931	HRIFA031249a : 2337
	HRIFA015486a : 1932	HRIFA031336a : 2338
	HRIFA015506a : 1933	HRIFA031350a : 2339
	HRIFA015536a : 1934	HRIFA031395a : 2340
40	HRIFA015547a : 1935	HRIFA031397a : 2341
	HRIFA015568a : 1936	HRIFA031438a : 2342
	HRIFA015671a : 1937	HRIFA031472a : 2343
	HRIFA015682a : 1938	HRIFA031510a : 2344
45	HRIFA015756a : 1939	HRIFA031672a : 2345
	HRIFA015764a : 1940	HRIFA031869a : 2346
	HRIFA015802a : 1941	HRIFA031871a : 2347
	HRIFA015811a : 1942	HRIFA031895a : 2348
	HRIFA015902a : 1943	HRIFA031935a : 2349
50	HRIFA015947a : 1944	HRIFA031986a : 2350
	HRIFA015995a : 1945	HRIFA032009a : 2351
	HRIFA016070a : 1946	HRIFA032011a : 2352
	HRIFA016129a : 1947	HRIFA032066a : 2353
	HRIFA016214a : 1948	HRIFA032067a : 2354
55	HRIFA016240a : 1949	HRIFA032070a : 2355

EP 1 130 094 A2

	HRIFA016255a : 1950	HRIFA032073a : 2356
	HRIFA016290a : 1951	HRIFA032079a : 2357
5	HRIFA016430a : 1952	HRIFA032097a : 2358
	HRIFA016599a : 1953	HRIFA032161a : 2359
	HRIFA016623a : 1954	HRIFA032186a : 2360
	HRIFA016639a : 1955	HRIFA032224a : 2361
10	HRIFA016654a : 1956	HRIFA032257a : 2362
	HRIFA016669a : 1957	HRIFA032271a : 2363
	HRIFA016758a : 1958	HRIFA032274a : 2364
	HRIFA016838a : 1959	HRIFA032275a : 2365
15	HRIFA016963a : 1960	HRIFA032360a : 2366
	HRIFA017031a : 1961	HRIFA032389a : 2367
	HRIFA017146a : 1962	HRIFA032433a : 2368
	HRIFA017190a : 1963	HRIFA032453a : 2369
20	HRIFA017257a : 1964	HRIFA032478a : 2370
	HRIFA017295a : 1965	HRIFA032506a : 2371
	HRIFA017312a : 1966	HRIFA032511a : 2372
	HRIFA017456a : 1967	HRIFA032530a : 2373
	HRIFA017457a : 1968	HRIFA032587a : 2374
25	HRIFA017509a : 1969	HRIFA032605a : 2375
	HRIFA017594a : 1970	HRIFA032642a : 2376
	HRIFA017643a : 1971	HRIFA032696a : 2377
	HRIFA017670a : 1972	HRIFA032730a : 2378
30	HRIFA017703a : 1973	HRIFA032820a : 2379
	HRIFA017729a : 1974	HRIFA032984a : 2380
	HRIFA017791a : 1975	HRIFA033349a : 2381
	HRIFA017801a : 1976	HRIFA033718a : 2382
35	HRIFA017818a : 1977	HRIFA034010a : 2383
	HRIFA017836a : 1978	

40

45

50

55

Table 369

The names of the internal sequences that are used in the selection of the clones from the representative sequences, and the corresponding SEQ IDs.

5

10

15

20

25

30

35

40

45

50

55

AA533598 : 2384	HRIFA036799a : 2463
AI051329 : 2385	HRIFA037138a : 2464
HRIFA000595a : 2386	HRIFA037776a : 2465
HRIFA000667a : 2387	HRIFA037838a : 2466
HRIFA000878a : 2388	HRIRA000001a : 2467
HRIFA001269a : 2389	HRIRA000041a : 2468
HRIFA001283a : 2390	HRIRA000058a : 2469
HRIFA002000a : 2391	HRIRA000260a : 2470
HRIFA002196a : 2392	HRIRA000490a : 2471
HRIFA003583a : 2393	HRIRA000522a : 2472
HRIFA005077a : 2394	HRIRA000553a : 2473
HRIFA005781a : 2395	HRIRA000563a : 2474
HRIFA006216a : 2396	HRIRA000640a : 2475
HRIFA006468a : 2397	HRIRA000725a : 2476
HRIFA006822a : 2398	HRIRA000998a : 2477
HRIFA007048a : 2399	HRIRA001053a : 2478
HRIFA007661a : 2400	HRIRA001314a : 2479

EP 1 130 094 A2

	HRIFA007777a : 2401	HRIRA001443a : 2480
	HRIFA007997a : 2402	HRIRA001473a : 2481
5	HRIFA008312a : 2403	HRIRA001648a : 2482
	HRIFA009250a : 2404	HRIRA001690a : 2483
	HRIFA009495a : 2405	HRIRA001726a : 2484
	HRIFA009607a : 2406	HRIRA001884a : 2485
	HRIFA009923a : 2407	HRIRA002098a : 2486
10	HRIFA009978a : 2408	HRIRA002100a : 2487
	HRIFA010730a : 2409	HRIRA002155a : 2488
	HRIFA011029a : 2410	HRIRA002307a : 2489
	HRIFA011416a : 2411	HRIRA002442a : 2490
	HRIFA011461a : 2412	HRIRA002446a : 2491
15	HRIFA012670a : 2413	HRIRA002479a : 2492
	HRIFA012717a : 2414	HRIRA002945a : 2493
	HRIFA012802a : 2415	HRIRA003028a : 2494
	HRIFA013357a : 2416	HRIRA003108a : 2495
	HRIFA013484a : 2417	HRIRA003139a : 2496
20	HRIFA015333a : 2418	HRIRA003819a : 2497
	HRIFA015375a : 2419	HRIRA004049a : 2498
	HRIFA015663a : 2420	HRIRA004286a : 2499
	HRIFA016287a : 2421	HRIRA004583a : 2500
	HRIFA016302a : 2422	HRIRA004691a : 2501
25	HRIFA016782a : 2423	HRIRA004783a : 2502
	HRIFA018555a : 2424	HRIRA005152a : 2503
	HRIFA019338a : 2425	HRIRA005221a : 2504
	HRIFA020315a : 2426	HRIRA005227a : 2505
	HRIFA020806a : 2427	HRIRA005305a : 2506
30	HRIFA022264a : 2428	HRIRA005563a : 2507
	HRIFA022923a : 2429	HRIRA006263a : 2508
	HRIFA023027a : 2430	HRIRA006324a : 2509
	HRIFA023218a : 2431	HRIRA006517a : 2510
35	HRIFA023363a : 2432	HRIRA006580a : 2511
	HRIFA023434a : 2433	HRIRA007665a : 2512
	HRIFA023444a : 2434	HRIRA007680a : 2513
	HRIFA023551a : 2435	HRIRA008129a : 2514
	HRIFA023558a : 2436	HRIRA008152a : 2515
40	HRIFA023641a : 2437	HRIRA008276a : 2516
	HRIFA023798a : 2438	HRIRA008329a : 2517
	HRIFA024330a : 2439	HRIRA008854a : 2518
	HRIFA024338a : 2440	HRIRA008896a : 2519
	HRIFA024384a : 2441	HRIRA008958a : 2520
45	HRIFA024644a : 2442	HRIRA009551a : 2521
	HRIFA025170a : 2443	HRIRA009828a : 2522
	HRIFA025496a : 2444	HRIRA010472a : 2523
	HRIFA025565a : 2445	HRIRA012442a : 2524
	HRIFA025651a : 2446	HRIRA012921a : 2525
50	HRIFA026224a : 2447	HRIRA013325a : 2526
	HRIFA026729a : 2448	HRIRA013644a : 2527
	HRIFA026925a : 2449	HRIRA013675a : 2528
	HRIFA028501a : 2450	HRIRA013702a : 2529
	HRIFA029454a : 2451	HRIRA013757a : 2530
55	HRIFA030181a : 2452	HRIRA013951a : 2531

EP 1 130 094 A2

5
10
15
20
25
30
35
40
45
50
55

HRIFA032701a : 2453	HRIRA014256a : 2532
HRIFA032801a : 2454	HRIRA014380a : 2533
HRIFA033384a : 2455	HRIRA015831a : 2534
HRIFA033682a : 2456	HRIRA015904a : 2535
HRIFA033930a : 2457	HRIRA016124a : 2536
HRIFA034817a : 2458	HRIRA017071a : 2537
HRIFA035409a : 2459	HRIRA018191a : 2538
HRIFA035542a : 2460	HRIRA020304a : 2539
HRIFA035577a : 2461	HRIRA000579a : 2540
HRIFA036630a : 2462	

[0286] The internal sequences include EST, HRIFA(the representative sequence of the 5'-end), and HRIRA (the representative sequence of the 3'-end).

Table 370

clone	the name of full-length nucleotide sequences	SEQ ID NO of the full-length nucleotide sequences	SEQ ID NO of the deduced amino acid sequences	SEQ ID NO
5	HEMBA1000006	C-HEMBA1000006	2547	2548
	nnnnnnnnnnnn	C-nnnnnnnnnnnn	nnnn	nnnn
15	HEMBA1000121	C-HEMBA1000121	2551	2552
	HEMBA1000128	C-HEMBA1000128	2553	2554
	HEMBA1000275	C-HEMBA1000275	2555	2556
	HEMBA1000300	C-HEMBA1000300	2557	
20	HEMBA1000349	C-HEMBA1000349	2558	2559
	HEMBA1000443	C-HEMBA1000443	2560	2561
	HEMBA1000590	C-HEMBA1000590	2562	2563
	HEMBA1000634	C-HEMBA1000634	2564	2565
25	HEMBA1000713	C-HEMBA1000713	2566	2567
	HEMBA1000745	C-HEMBA1000745	2568	2569
	HEMBA1000907	C-HEMBA1000907	2570	2571
	HEMBA1000940	C-HEMBA1000940	2572	2573
	HEMBA1000962	C-HEMBA1000962	2574	2575
30	HEMBA1001221	C-HEMBA1001221	2576	2577
	HEMBA1001228	C-HEMBA1001228	2578	2579
	HEMBA1001297	C-HEMBA1001297	2580	
	HEMBA1001390	C-HEMBA1001390	2581	2582
35	HEMBA1001563	C-HEMBA1001563	2583	
	HEMBA1001621	C-HEMBA1001621	2584	2585
	nnnnnnnnnnnn	C-nnnnnnnnnnnn	nnnn	nnnn
	HEMBA1001878	C-HEMBA1001878	2588	2589
40	HEMBA1002131	C-HEMBA1002131	2590	2591
	HEMBA1002163	C-HEMBA1002163	2592	2593
	HEMBA1002164	C-HEMBA1002164	2594	2595
	HEMBA1002167	C-HEMBA1002167	2596	2597
45	HEMBA1002178	C-HEMBA1002178	2598	2599
	nnnnnnnnnnnn	C-nnnnnnnnnnnn	nnnn	nnnn

50

55

EP 1 130 094 A2

	HEMBA1002195	C-HEMBA1002195	2602	2603
	HEMBA1002227	C-HEMBA1002227	2604	2605
	HEMBA1002239	C-HEMBA1002239	2606	
5	HEMBA1002316	C-HEMBA1002316	2607	2608
	HEMBA1002420	C-HEMBA1002420	2609	2610
	HEMBA1002421	C-HEMBA1002421	2611	2612
	HEMBA1002524	C-HEMBA1002524	2613	2614
	HEMBA1002551	C-HEMBA1002551	2615	2616
10	HEMBA1002767	C-HEMBA1002767	2617	2618
	HEMBA1002992	C-HEMBA1002992	2619	2620
	HEMBA1003047	C-HEMBA1003047	2621	2622
	HEMBA1003072	C-HEMBA1003072	2623	2624
15	HEMBA1003101	C-HEMBA1003101	2625	2626
	HEMBA1003230	C-HEMBA1003230	2627	2628
	HEMBA1003294	C-HEMBA1003294	2629	
	HEMBA1003315	C-HEMBA1003315	2630	2631
	HEMBA1003392	C-HEMBA1003392	2632	2633
20	HEMBA1003399	C-HEMBA1003399	2634	2635
	HEMBA1003487	C-HEMBA1003487	2636	2637
	HEMBA1003530	C-HEMBA1003530	2638	2639
	HEMBA1003602	C-HEMBA1003602	2640	2641
	HEMBA1003732	C-HEMBA1003732	2642	2643
25	HEMBA1003945	C-HEMBA1003945	2644	2645
	HEMBA1004110	C-HEMBA1004110	2646	2647
	HEMBA1004250	C-HEMBA1004250	2648	2649
	HEMBA1004391	C-HEMBA1004391	2650	2651
	HEMBA1004444	C-HEMBA1004444	2652	2653
30	HEMBA1004454	C-HEMBA1004454	2654	2655
	HEMBA1004505	C-HEMBA1004505	2656	2657
	HEMBA1004797	C-HEMBA1004797	2658	2659
	HEMBA1004982	C-HEMBA1004982	2660	2661
35	HEMBA1005070	C-HEMBA1005070	2662	2663
	HEMBA1005084	C-HEMBA1005084	2664	2665
	HEMBA1005145	C-HEMBA1005145	2666	2667
	HEMBA1005430	C-HEMBA1005430	2668	2669
	HEMBA1005449	C-HEMBA1005449	2670	2671
40	HEMBA1005489	C-HEMBA1005489	2672	2673
	HEMBA1005522	C-HEMBA1005522	2674	2675
	HEMBA1005545	C-HEMBA1005545	2676	2677
	HEMBA1005698	C-HEMBA1005698	2678	2679
	HEMBA1005913	C-HEMBA1005913	2680	
45	HEMBA1005929	C-HEMBA1005929	2681	2682
	HEMBA1005945	C-HEMBA1005945	2683	2684
	HEMBA1006016	C-HEMBA1006016	2685	
	HEMBA1006171	C-HEMBA1006171	2686	2687
	HEMBA1006299	C-HEMBA1006299	2688	2689
50	HEMBA1006311	C-HEMBA1006311	2690	2691
	HEMBA1006335	C-HEMBA1006335	2692	2693
	HEMBA1006430	C-HEMBA1006430	2694	2695
	HEMBA1006482	C-HEMBA1006482	2696	2697
	HEMBA1006572	C-HEMBA1006572	2698	2699
55	HEMBA1006707	C-HEMBA1006707	2700	2701

EP 1 130 094 A2

	HEMBA1006724	C-HEMBA1006724	2702	2703
	HEMBA1006902	C-HEMBA1006902	2704	2705
	HEMBA1006916	C-HEMBA1006916	2706	2707
5	HEMBA1006960	C-HEMBA1006960	2708	2709
	HEMBA1007013	C-HEMBA1007013	2710	2711
	HEMBA1007057	C-HEMBA1007057	2712	2713
	HEMBA1007241	C-HEMBA1007241	2714	
	HEMBA1007291	C-HEMBA1007291	2715	2716
10	HEMBA1007332	C-HEMBA1007332	2717	
	HEMBB1000276	C-HEMBB1000276	2718	
	HEMBB1000447	C-HEMBB1000447	2719	2720
	HEMBB1000642	C-HEMBB1000642	2721	
15	HEMBB1000668	C-HEMBB1000668	2722	2723
	HEMBB1000679	C-HEMBB1000679	2724	2725
	HEMBB1000881	C-HEMBB1000881	2726	2727
	HEMBB1000905	C-HEMBB1000905	2728	2729
	HEMBB1001026	C-HEMBB1001026	2730	2731
20	HEMBB1001048	C-HEMBB1001048	2732	2733
	HEMBB1001200	C-HEMBB1001200	2734	
	HEMBB1001407	C-HEMBB1001407	2735	2736
	HEMBB1001530	C-HEMBB1001530	2737	2738
	HEMBB1001573	C-HEMBB1001573	2739	2740
25	nnnnnnnnnnnn C-nnnnnnnnnnnnn nnnn	nnnn	nnnn	
	HEMBB1001847	C-HEMBB1001847	2743	2744
	HEMBB1001978	C-HEMBB1001978	2745	2746
	HEMBB1002162	C-HEMBB1002162	2747	2748
	HEMBB1002228	C-HEMBB1002228	2749	
30	HEMBB1002245	C-HEMBB1002245	2750	2751
	HEMBB1002427	C-HEMBB1002427	2752	2753
	HEMBB1002465	C-HEMBB1002465	2754	2755
	HEMBB1002663	C-HEMBB1002663	2756	2757
35	HEMBB1002693	C-HEMBB1002693	2758	2759
	MAMMA1000046	C-MAMMA1000046	2760	
	MAMMA1000118	C-MAMMA1000118	2761	2762
	nnnnnnnnnnnn C-nnnnnnnnnnnnn nnnn	nnnn	nnnn	
	MAMMA1000449	C-MAMMA1000449	2765	
40	MAMMA1000457	C-MAMMA1000457	2766	2767
	MAMMA1000591	C-MAMMA1000591	2768	2769
	MAMMA1000681	C-MAMMA1000681	2770	2771
	MAMMA1001043	C-MAMMA1001043	2772	2773
	MAMMA1001893	C-MAMMA1001893	2774	2775
45	NT2RM2000241	C-NT2RM2000241	2776	2777
	NT2RM2000306	C-NT2RM2000306	2778	2779
	NT2RM2000410	C-NT2RM2000410	2780	2781
	NT2RM2000423	C-NT2RM2000423	2782	2783
	NT2RM2000497	C-NT2RM2000497	2784	2785
50	NT2RM2000514	C-NT2RM2000514	2786	2787
	NT2RM2000622	C-NT2RM2000622	2788	2789
	NT2RM2001126	C-NT2RM2001126	2790	2791
	NT2RM2001902	C-NT2RM2001902	2792	2793
	NT2RM2001939	C-NT2RM2001939	2794	2795
55	NT2RM2001941	C-NT2RM2001941	2796	2797

EP 1 130 094 A2

	NT2RM4000198	C-NT2RM4000198	2798	2799
	NT2RM4000284	C-NT2RM4000284	2800	2801
	NT2RM4000295	C-NT2RM4000295	2802	2803
5	NT2RM4000326	C-NT2RM4000326	2804	2805
	NT2RM4000444	C-NT2RM4000444	2806	2807
	NT2RM4000587	C-NT2RM4000587	2808	2809
	NT2RM4000648	C-NT2RM4000648	2810	2811
10	NT2RM4000997	C-NT2RM4000997	2812	2813
	NT2RM4001321	C-NT2RM4001321	2814	2815
	NT2RM4001325	C-NT2RM4001325	2816	2817
	NT2RM4001735	C-NT2RM4001735	2818	2819
	NT2RM4002352	C-NT2RM4002352	2820	2821
15	NT2RP1000002	C-NT2RP1000002	2822	2823
	NT2RP1000050	C-NT2RP1000050	2824	2825
	NT2RP1000181	C-NT2RP1000181	2826	2827
	NT2RP1000261	C-NT2RP1000261	2828	2829
	NT2RP1000300	C-NT2RP1000300	2830	2831
20	NT2RP1000325	C-NT2RP1000325	2832	2833
	NT2RP1000448	C-NT2RP1000448	2834	2835
	NT2RP1000551	C-NT2RP1000551	2836	2837
	NT2RP1000579	C-NT2RP1000579	2838	2839
25	NT2RP1000613	C-NT2RP1000613	2840	2841
	NT2RP1000903	C-NT2RP1000903	2842	2843
	NT2RP1000981	C-NT2RP1000981	2844	2845
	NT2RP1001004	C-NT2RP1001004	2846	2847
	NT2RP1001020	C-NT2RP1001020	2848	2849
30	NT2RP1001563	C-NT2RP1001563	2850	2851
	NT2RP2000394	C-NT2RP2000394	2852	2853
	NT2RP2000479	C-NT2RP2000479	2854	2855
	NT2RP2000514	C-NT2RP2000514	2856	2857
	NT2RP2000533	C-NT2RP2000533	2858	2859
35	NT2RP2000649	C-NT2RP2000649	2860	2861
	NT2RP2000663	C-NT2RP2000663	2862	2863
	NT2RP2000694	C-NT2RP2000694	2864	2865
	NT2RP2000903	C-NT2RP2000903	2866	2867
40	NT2RP2001480	C-NT2RP2001480	2868	2869
	NT2RP2001495	C-NT2RP2001495	2870	2871
	NT2RP2001514	C-NT2RP2001514	2872	2873
	NT2RP2001529	C-NT2RP2001529	2874	2875
	NT2RP2001769	C-NT2RP2001769	2876	2877
45	NT2RP2001878	C-NT2RP2001878	2878	2879
	NT2RP2001903	C-NT2RP2001903	2880	2881
	NT2RP2001915	C-NT2RP2001915	2882	2883
	NT2RP2001956	C-NT2RP2001956	2884	2885
	NT2RP2002063	C-NT2RP2002063	2886	2887
50	NT2RP2002188	C-NT2RP2002188	2888	2889
	NT2RP2002232	C-NT2RP2002232	2890	2891
	NT2RP2002304	C-NT2RP2002304	2892	2893
	NT2RP2002409	C-NT2RP2002409	2894	2895
	NT2RP2002510	C-NT2RP2002510	2896	2897
55	NT2RP2002527	C-NT2RP2002527	2898	2899
	NT2RP2002533	C-NT2RP2002533	2900	2901

EP 1 130 094 A2

	NT2RP2002564	C-NT2RP2002564	2902	2903
	NT2RP2002674	C-NT2RP2002674	2904	2905
	NT2RP2002721	C-NT2RP2002721	2906	2907
5	NT2RP2002824	C-NT2RP2002824	2908	2909
	NT2RP2002942	C-NT2RP2002942	2910	2911
	NT2RP2002974	C-NT2RP2002974	2912	2913
	NT2RP2002976	C-NT2RP2002976	2914	2915
	NT2RP2003042	C-NT2RP2003042	2916	2917
10	NT2RP2003179	C-NT2RP2003179	2918	2919
	NT2RP2003210	C-NT2RP2003210	2920	2921
	NT2RP2003369	C-NT2RP2003369	2922	2923
	NT2RP2003383	C-NT2RP2003383	2924	2925
15	NT2RP2003469	C-NT2RP2003469	2926	2927
	NT2RP2003545	C-NT2RP2003545	2928	2929
	NT2RP2003593	C-NT2RP2003593	2930	2931
	NT2RP2003599	C-NT2RP2003599	2932	2933
	NT2RP2003655	C-NT2RP2003655	2934	2935
20	NT2RP2003931	C-NT2RP2003931	2936	2937
	NT2RP2004141	C-NT2RP2004141	2938	2939
	NT2RP2004179	C-NT2RP2004179	2940	2941
	NT2RP2004205	C-NT2RP2004205	2942	2943
	NT2RP2004447	C-NT2RP2004447	2944	2945
25	NT2RP2004495	C-NT2RP2004495	2946	2947
	NT2RP2004524	C-NT2RP2004524	2948	2949
	NT2RP2004556	C-NT2RP2004556	2950	2951
	NT2RP2004606	C-NT2RP2004606	2952	2953
	NT2RP2004648	C-NT2RP2004648	2954	2955
30	NT2RP2004670	C-NT2RP2004670	2956	2957
	NT2RP2004794	C-NT2RP2004794	2958	2959
	NT2RP2004837	C-NT2RP2004837	2960	2961
	NT2RP2004847	C-NT2RP2004847	2962	2963
35	nnnnnnnnnnnn	C-nnnnnnnnnnnn	nnnn	nnnn
	NT2RP2005027	C-NT2RP2005027	2966	
	NT2RP2005163	C-NT2RP2005163	2967	2968
	NT2RP2005181	C-NT2RP2005181	2969	2970
	NT2RP2005247	C-NT2RP2005247	2971	2972
40	NT2RP2005425	C-NT2RP2005425	2973	2974
	NT2RP2005463	C-NT2RP2005463	2975	2976
	NT2RP2005514	C-NT2RP2005514	2977	2978
	NT2RP2005541	C-NT2RP2005541	2979	2980
	NT2RP2005632	C-NT2RP2005632	2981	2982
45	NT2RP2005878	C-NT2RP2005878	2983	2984
	NT2RP2005883	C-NT2RP2005883	2985	2986
	NT2RP2005887	C-NT2RP2005887	2987	2988
	NT2RP2005941	C-NT2RP2005941	2989	2990
	NT2RP2005994	C-NT2RP2005994	2991	2992
50	NT2RP2006042	C-NT2RP2006042	2993	2994
	NT2RP2006269	C-NT2RP2006269	2995	2996
	NT2RP2006512	C-NT2RP2006512	2997	2998
	NT2RP3000059	C-NT2RP3000059	2999	3000
	NT2RP3000063	C-NT2RP3000063	3001	3002
55	NT2RP3000125	C-NT2RP3000125	3003	3004

EP 1 130 094 A2

[illegible]

EP 1 130 094 A2

	NT2RP3004202	C-NT2RP3004202	3109	3110
	NT2RP3004309	C-NT2RP3004309	3111	3112
	NT2RP3004321	C-NT2RP3004321	3113	3114
5	NT2RP3004355	C-NT2RP3004355	3115	3116
	NT2RP3004374	C-NT2RP3004374	3117	3118
	NT2RP3004406	C-NT2RP3004406	3119	3120
	NT2RP3004552	C-NT2RP3004552	3121	3122
	NT2RP3004557	C-NT2RP3004557	3123	3124
10	NT2RP3004625	C-NT2RP3004625	3125	3126
	NT2RP3004640	C-NT2RP3004640	3127	3128
	NT2RP3004647	C-NT2RP3004647	3129	3130
	NT2RP4000108	C-NT2RP4000108	3131	3132
	NT2RP4000962	C-NT2RP4000962	3133	3134
15	NT2RP4001009	C-NT2RP4001009	3135	3136
	NT2RP4001467	C-NT2RP4001467	3137	3138
	OVARC1000090	C-OVARC1000090	3139	
	OVARC1000105	C-OVARC1000105	3140	3141
20	OVARC1000137	C-OVARC1000137	3142	
	OVARC1000208	C-OVARC1000208	3143	
	OVARC1000255	C-OVARC1000255	3144	3145
	OVARC1000275	C-OVARC1000275	3146	3147
	OVARC1000298	C-OVARC1000298	3148	3149
25	OVARC1000410	C-OVARC1000410	3150	3151
	OVARC1000439	C-OVARC1000439	3152	3153
	OVARC1000467	C-OVARC1000467	3154	3155
	OVARC1000529	C-OVARC1000529	3156	3157
	OVARC1000775	C-OVARC1000775	3158	3159
30	nnnnnnnnnnnnnn C-nnnnnnnnnnnnn	nnnn	nnnn	
	OVARC1000811	C-OVARC1000811	3162	3163
	OVARC1000853	C-OVARC1000853	3164	3165
	OVARC1000916	C-OVARC1000916	3166	3167
	OVARC1000956	C-OVARC1000956	3168	3169
35	OVARC1001030	C-OVARC1001030	3170	3171
	OVARC1001049	C-OVARC1001049	3172	3173
	OVARC1001086	C-OVARC1001086	3174	3175
	OVARC1001132	C-OVARC1001132	3176	3177
40	OVARC1001163	C-OVARC1001163	3178	3179
	OVARC1001222	C-OVARC1001222	3180	
	OVARC1001338	C-OVARC1001338	3181	3182
	OVARC1001569	C-OVARC1001569	3183	3184
	OVARC1001596	C-OVARC1001596	3185	3186
45	OVARC1001725	C-OVARC1001725	3187	
	OVARC1001727	C-OVARC1001727	3188	
	OVARC1001807	C-OVARC1001807	3189	3190
	OVARC1001991	C-OVARC1001991	3191	3192
	OVARC1002058	C-OVARC1002058	3193	3194
50	OVARC1002178	C-OVARC1002178	3195	
	PLACE1000033	C-PLACE1000033	3196	3197
	PLACE1000231	C-PLACE1000231	3198	3199
	PLACE1000258	C-PLACE1000258	3200	3201
	PLACE1000442	C-PLACE1000442	3202	3203
55	PLACE1000560	C-PLACE1000560	3204	3205

EP 1 130 094 A2

	PLACE1000740	C-PLACE1000740	3206	3207
	PLACE1000912	C-PLACE1000912	3208	3209
	PLACE1000914	C-PLACE1000914	3210	3211
5	PLACE1000927	C-PLACE1000927	3212	3213
	PLACE1001016	C-PLACE1001016	3214	3215
	nnnnnnnnnnnn	C-nnnnnnnnnnnnn	nnnn	nnnn
	PLACE1001100	C-PLACE1001100	3218	
	PLACE1001114	C-PLACE1001114	3219	3220
10	PLACE1001123	C-PLACE1001123	3221	
	PLACE1001183	C-PLACE1001183	3222	3223
	PLACE1001229	C-PLACE1001229	3224	3225
	PLACE1001231	C-PLACE1001231	3226	3227
15	nnnnnnnnnnnn	C-nnnnnnnnnnnnn	nnnn	nnnn
	PLACE1001340	C-PLACE1001340	3230	3231
	PLACE1001401	C-PLACE1001401	3232	3233
	PLACE1001407	C-PLACE1001407	3234	3235
	PLACE1001464	C-PLACE1001464	3236	3237
20	PLACE1001500	C-PLACE1001500	3238	3239
	PLACE1001516	C-PLACE1001516	3240	3241
	PLACE1001536	C-PLACE1001536	3242	3243
	PLACE1001564	C-PLACE1001564	3244	3245
	PLACE1001655	C-PLACE1001655	3246	3247
25	nnnnnnnnnnnn	C-nnnnnnnnnnnnn	nnnn	nnnn
	PLACE1001788	C-PLACE1001788	3250	3251
	PLACE1001795	C-PLACE1001795	3252	3253
	PLACE1001836	C-PLACE1001836	3254	3255
	PLACE1001918	C-PLACE1001918	3256	3257
30	PLACE1001949	C-PLACE1001949	3258	3259
	PLACE1002080	C-PLACE1002080	3260	3261
	PLACE1002095	C-PLACE1002095	3262	3263
	PLACE1002153	C-PLACE1002153	3264	3265
35	PLACE1002329	C-PLACE1002329	3266	3267
	PLACE1002355	C-PLACE1002355	3268	3269
	PLACE1002374	C-PLACE1002374	3270	3271
	PLACE1002518	C-PLACE1002518	3272	3273
	PLACE1002547	C-PLACE1002547	3274	3275
40	PLACE1002726	C-PLACE1002726	3276	3277
	PLACE1002905	C-PLACE1002905	3278	3279
	PLACE1002911	C-PLACE1002911	3280	3281
	PLACE1002967	C-PLACE1002967	3282	3283
	PLACE1003135	C-PLACE1003135	3284	3285
45	PLACE1003163	C-PLACE1003163	3286	3287
	PLACE1003428	C-PLACE1003428	3288	3289
	PLACE1003438	C-PLACE1003438	3290	3291
	PLACE1003460	C-PLACE1003460	3292	3293
	PLACE1003529	C-PLACE1003529	3294	3295
50	PLACE1003573	C-PLACE1003573	3296	3297
	PLACE1003598	C-PLACE1003598	3298	3299
	PLACE1003644	C-PLACE1003644	3300	
	PLACE1003737	C-PLACE1003737	3301	3302
	PLACE1003772	C-PLACE1003772	3303	3304
55	PLACE1003839	C-PLACE1003839	3305	3306

EP 1 130 094 A2

5

EP 1 130 094 A2

	PLACE1007591	C-PLACE1007591	3410	
	PLACE1007626	C-PLACE1007626	3411	3412
	PLACE1007702	C-PLACE1007702	3413	3414
5	PLACE1007845	C-PLACE1007845	3415	3416
	PLACE1007881	C-PLACE1007881	3417	3418
	PLACE1008297	C-PLACE1008297	3419	3420
	nnnnnnnnnnnn	C-nnnnnnnnnnnnn	nnnn	nnnn
	PLACE1008469	C-PLACE1008469	3423	3424
10	PLACE1008549	C-PLACE1008549	3425	3426
	PLACE1008657	C-PLACE1008657	3427	3428
	PLACE1008716	C-PLACE1008716	3429	3430
	PLACE1008984	C-PLACE1008984	3431	3432
15	PLACE1008985	C-PLACE1008985	3433	3434
	PLACE1009067	C-PLACE1009067	3435	3436
	PLACE1009196	C-PLACE1009196	3437	3438
	PLACE1009279	C-PLACE1009279	3439	3440
	PLACE1009527	C-PLACE1009527	3441	3442
20	PLACE1009546	C-PLACE1009546	3443	3444
	PLACE1009600	C-PLACE1009600	3445	3446
	PLACE1009735	C-PLACE1009735	3447	3448
	PLACE1009982	C-PLACE1009982	3449	3450
	PLACE1010078	C-PLACE1010078	3451	3452
25	PLACE1010081	C-PLACE1010081	3453	3454
	PLACE1010251	C-PLACE1010251	3455	3456
	PLACE1010784	C-PLACE1010784	3457	3458
	PLACE1010827	C-PLACE1010827	3459	3460
	PLACE1010968	C-PLACE1010968	3461	3462
30	PLACE1011045	C-PLACE1011045	3463	3464
	PLACE1011116	C-PLACE1011116	3465	3466
	PLACE1011236	C-PLACE1011236	3467	3468
	PLACE1011407	C-PLACE1011407	3469	3470
35	PLACE1011516	C-PLACE1011516	3471	3472
	PLACE1011708	C-PLACE1011708	3473	3474
	PLACE1011824	C-PLACE1011824	3475	3476
	PLACE1011978	C-PLACE1011978	3477	3478
	PLACE2000118	C-PLACE2000118	3479	3480
40	PLACE2000219	C-PLACE2000219	3481	
	SKNMC1000004	C-SKNMC1000004	3482	3483
	THYRO1000036	C-THYRO1000036	3484	3485
	THYRO1000061	C-THYRO1000061	3486	3487
	THYRO1000099	C-THYRO1000099	3488	3489
45	THYRO1000196	C-THYRO1000196	3490	3491
	THYRO1000400	C-THYRO1000400	3492	3493
	THYRO1000580	C-THYRO1000580	3494	3495
	THYRO1000584	C-THYRO1000584	3496	3497
	THYRO1000678	C-THYRO1000678	3498	3499
50	THYRO1000795	C-THYRO1000795	3500	3501
	THYRO1000846	C-THYRO1000846	3502	3503
	THYRO1000866	C-THYRO1000866	3504	3505
	THYRO1000956	C-THYRO1000956	3506	3507
	THYRO1000999	C-THYRO1000999	3508	
55	THYRO1001063	C-THYRO1001063	3509	3510

EP 1 130 094 A2

	THYRO1001071	C-THYRO1001071	3511	3512
	THYRO1001102	C-THYRO1001102	3513	3514
	THYRO1001113	C-THYRO1001113	3515	3516
5	THYRO1001128	C-THYRO1001128	3517	3518
	THYRO1001205	C-THYRO1001205	3519	3520
	THYRO1001237	C-THYRO1001237	3521	3522
	THYRO1001266	C-THYRO1001266	3523	3524
10	THYRO1001327	C-THYRO1001327	3525	3526
	THYRO1001456	C-THYRO1001456	3527	3528
	THYRO1001457	C-THYRO1001457	3529	3530
	THYRO1001471	C-THYRO1001471	3531	3532
	THYRO1001478	C-THYRO1001478	3533	3534
15	THYRO1001495	C-THYRO1001495	3535	3536
	THYRO1001523	C-THYRO1001523	3537	3538
	THYRO1001529	C-THYRO1001529	3539	3540
	THYRO1001700	C-THYRO1001700	3541	3542
	THYRO1001702	C-THYRO1001702	3543	3544
20	THYRO1001725	C-THYRO1001725	3545	3546
	THYRO1001803	C-THYRO1001803	3547	3548
	Y79AA1000127	C-Y79AA1000127	3549	3550
	Y79AA1000207	C-Y79AA1000207	3551	3552
25	Y79AA1000226	C-Y79AA1000226	3553	3554
	Y79AA1000270	C-Y79AA1000270	3555	3556
	Y79AA1000426	C-Y79AA1000426	3557	3558
	Y79AA1000521	C-Y79AA1000521	3559	3560
	Y79AA1000776	C-Y79AA1000776	3561	3562
30	Y79AA1000777	C-Y79AA1000777	3563	3564
	nnnnnnnnnnnnnn C-nnnnnnnnnnnnnn nnnnn	nnnnn	nnnnn	
	Y79AA1000876	C-Y79AA1000876	3567	3568
	Y79AA1000959	C-Y79AA1000959	3569	3570
	Y79AA1000967	C-Y79AA1000967	3571	3572
35	Y79AA1001013	C-Y79AA1001013	3573	3574
	Y79AA1001056	C-Y79AA1001056	3575	3576
	Y79AA1001062	C-Y79AA1001062	3577	3578
	Y79AA1001090	C-Y79AA1001090	3579	3580
	Y79AA1001264	C-Y79AA1001264	3581	3582
40	Y79AA1001272	C-Y79AA1001272	3583	3584
	Y79AA1001328	C-Y79AA1001328	3585	3586
	Y79AA1001430	C-Y79AA1001430	3587	3588
	Y79AA1002022	C-Y79AA1002022	3589	3590
45	BNGH41000020	C-BNGH41000020	3595	3596
	BNGH41000091	C-BNGH41000091	3597	3598
	HEMBA1000462	C-HEMBA1000462	3599	3600
	HEMBA1000477	C-HEMBA1000477	3601	3602
	HEMBA1000671	C-HEMBA1000671	3603	3604
50	HEMBA1000732	C-HEMBA1000732	3605	3606
	HEMBA1000835	C-HEMBA1000835	3607	3608
	HEMBA1000875	C-HEMBA1000875	3609	
	HEMBA1001184	C-HEMBA1001184	3610	
	HEMBA1001272	C-HEMBA1001272	3611	
55	HEMBA1001296	C-HEMBA1001296	3612	

EP 1 130 094 A2

	HEMBA1002048	C-HEMBA1002048	3613	3614
	HEMBA1002985	C-HEMBA1002985	3615	3616
	HEMBA1003120	C-HEMBA1003120	3617	3618
5	HEMBA1003497	C-HEMBA1003497	3619	3620
	HEMBA1004007	C-HEMBA1004007	3621	
	HEMBA1004085	C-HEMBA1004085	3622	3623
	HEMBA1004785	C-HEMBA1004785	3624	3625
	HEMBA1004952	C-HEMBA1004952	3626	
10	HEMBA1004971	C-HEMBA1004971	3627	
	HEMBA1005230	C-HEMBA1005230	3628	
	HEMBA1005246	C-HEMBA1005246	3629	3630
	HEMBA1005267	C-HEMBA1005267	3631	3632
15	HEMBA1006276	C-HEMBA1006276	3633	
	HEMBA1006357	C-HEMBA1006357	3634	3635
	HEMBA1006517	C-HEMBA1006517	3636	3637
	HEMBA1006544	C-HEMBA1006544	3638	3639
	HEMBA1006749	C-HEMBA1006749	3640	3641
20	HEMBA1006770	C-HEMBA1006770	3642	3643
	HEMBA1006912	C-HEMBA1006912	3644	
	HEMBA1007063	C-HEMBA1007063	3645	3646
	HEMBB1000106	C-HEMBB1000106	3647	3648
	HEMBB1000407	C-HEMBB1000407	3649	3650
25	HEMBB1000542	C-HEMBB1000542	3651	3652
	HEMBB1001547	C-HEMBB1001547	3653	3654
	HEMBB1001959	C-HEMBB1001959	3655	3656
	HEMBB1002039	C-HEMBB1002039	3657	
	HEMBB1002041	C-HEMBB1002041	3658	3659
30	HEMBB1002051	C-HEMBB1002051	3660	3661
	HEMBB1002120	C-HEMBB1002120	3662	3663
	HEMBB1002302	C-HEMBB1002302	3664	3665
	HEMBB1002661	C-HEMBB1002661	3666	3667
35	MAMMA1000106	C-MAMMA1000106	3668	3669
	MAMMA1000141	C-MAMMA1000141	3670	3671
	MAMMA1000204	C-MAMMA1000204	3672	3673
	MAMMA1000226	C-MAMMA1000226	3674	3675
	MAMMA1000403	C-MAMMA1000403	3676	3677
40	MAMMA1000473	C-MAMMA1000473	3678	3679
	MAMMA1000496	C-MAMMA1000496	3680	3681
	MAMMA1000528	C-MAMMA1000528	3682	
	MAMMA1000614	C-MAMMA1000614	3683	3684
	MAMMA1000652	C-MAMMA1000652	3685	
45	MAMMA1000706	C-MAMMA1000706	3686	3687
	MAMMA1000788	C-MAMMA1000788	3688	3689
	MAMMA1000810	C-MAMMA1000810	3690	3691
	MAMMA1000814	C-MAMMA1000814	3692	3693
	MAMMA1000881	C-MAMMA1000881	3694	3695
50	MAMMA1000986	C-MAMMA1000986	3696	3697
	MAMMA1000994	C-MAMMA1000994	3698	3699
	MAMMA1001141	C-MAMMA1001141	3700	3701
	MAMMA1001150	C-MAMMA1001150	3702	3703
	MAMMA1001237	C-MAMMA1001237	3704	3705
55	MAMMA1001284	C-MAMMA1001284	3706	3707

EP 1 130 094 A2

	MAMMA1001310	C-MAMMA1001310	3708	3709
	MAMMA1001344	C-MAMMA1001344	3710	3711
	MAMMA1001418	C-MAMMA1001418	3712	3713
5	MAMMA1001532	C-MAMMA1001532	3714	3715
	MAMMA1001615	C-MAMMA1001615	3716	3717
	MAMMA1001623	C-MAMMA1001623	3718	3719
	MAMMA1001634	C-MAMMA1001634	3720	3721
10	MAMMA1001957	C-MAMMA1001957	3722	3723
	MAMMA1001978	C-MAMMA1001978	3724	3725
	MAMMA1002070	C-MAMMA1002070	3726	3727
	MAMMA1002080	C-MAMMA1002080	3728	3729
	MAMMA1002087	C-MAMMA1002087	3730	3731
15	MAMMA1002095	C-MAMMA1002095	3732	3733
	MAMMA1002128	C-MAMMA1002128	3734	3735
	MAMMA1002142	C-MAMMA1002142	3736	3737
	MAMMA1002165	C-MAMMA1002165	3738	3739
	MAMMA1002205	C-MAMMA1002205	3740	
20	MAMMA1002234	C-MAMMA1002234	3741	3742
	MAMMA1002586	C-MAMMA1002586	3743	3744
	MAMMA1002633	C-MAMMA1002633	3745	3746
	MAMMA1003126	C-MAMMA1003126	3747	3748
25	NT2RM1000580	C-NT2RM1000580	3749	3750
	NT2RM1000858	C-NT2RM1000858	3751	3752
	NT2RM2000565	C-NT2RM2000565	3753	3754
	NT2RM2000582	C-NT2RM2000582	3755	
	NT2RM2000589	C-NT2RM2000589	3756	3757
30	NT2RM2000632	C-NT2RM2000632	3758	3759
	NT2RM2000773	C-NT2RM2000773	3760	
	NT2RM2001558	C-NT2RM2001558	3761	3762
	NT2RM2001626	C-NT2RM2001626	3763	3764
	NT2RM2001643	C-NT2RM2001643	3765	3766
35	NT2RM2001738	C-NT2RM2001738	3767	3768
	NT2RM2001792	C-NT2RM2001792	3769	3770
	NT2RM2001818	C-NT2RM2001818	3771	3772
	NT2RM4000100	C-NT2RM4000100	3773	3774
	NT2RM4000115	C-NT2RM4000115	3775	3776
40	NT2RM4000417	C-NT2RM4000417	3777	3778
	NT2RM4000593	C-NT2RM4000593	3779	3780
	NT2RM4000761	C-NT2RM4000761	3781	3782
	NT2RM4000965	C-NT2RM4000965	3783	3784
45	NT2RM4001377	C-NT2RM4001377	3785	3786
	NT2RM4001768	C-NT2RM4001768	3787	3788
	NT2RM4001843	C-NT2RM4001843	3789	3790
	NT2RP1000239	C-NT2RP1000239	3791	3792
	NT2RP1000465	C-NT2RP1000465	3793	3794
50	NT2RP1000468	C-NT2RP1000468	3795	3796
	NT2RP1000679	C-NT2RP1000679	3797	3798
	NT2RP1000740	C-NT2RP1000740	3799	3800
	NT2RP1001031	C-NT2RP1001031	3801	3802
	NT2RP2000178	C-NT2RP2000178	3803	3804
55	NT2RP2000240	C-NT2RP2000240	3805	
	NT2RP2000447	C-NT2RP2000447	3806	3807

EP 1 130 094 A2

	NT2RP2000610	C-NT2RP2000610	3808	3809
	NT2RP2000616	C-NT2RP2000616	3810	3811
5	NT2RP2000712	C-NT2RP2000712	3812	3813
	NT2RP2000739	C-NT2RP2000739	3814	3815
	NT2RP2000818	C-NT2RP2000818	3816	3817
	NT2RP2001200	C-NT2RP2001200	3818	3819
	NT2RP2001223	C-NT2RP2001223	3820	3821
10	NT2RP2001276	C-NT2RP2001276	3822	3823
	NT2RP2001388	C-NT2RP2001388	3824	3825
	NT2RP2001469	C-NT2RP2001469	3826	3827
	NT2RP2001562	C-NT2RP2001562	3828	3829
	NT2RP2001662	C-NT2RP2001662	3830	3831
15	NT2RP2001755	C-NT2RP2001755	3832	3833
	NT2RP2001817	C-NT2RP2001817	3834	3835
	NT2RP2001948	C-NT2RP2001948	3836	3837
	NT2RP2002015	C-NT2RP2002015	3838	3839
	NT2RP2003390	C-NT2RP2003390	3840	3841
20	NT2RP2003664	C-NT2RP2003664	3842	3843
	NT2RP2003940	C-NT2RP2003940	3844	3845
	NT2RP2004069	C-NT2RP2004069	3846	3847
	NT2RP2004108	C-NT2RP2004108	3848	3849
25	nnnnnnnnnnnn C-nnnnnnnnnnnnn		3850	3851
	NT2RP2005069	C-NT2RP2005069	3852	3853
	NT2RP2005378	C-NT2RP2005378	3854	3855
	NT2RP2005391	C-NT2RP2005391	3856	3857
	NT2RP2005597	C-NT2RP2005597	3858	3859
30	NT2RP2005666	C-NT2RP2005666	3860	3861
	NT2RP2006004	C-NT2RP2006004	3862	3863
	NT2RP2006092	C-NT2RP2006092	3864	3865
	NT2RP2006134	C-NT2RP2006134	3866	3867
	NT2RP3000011	C-NT2RP3000011	3868	3869
35	NT2RP3000022	C-NT2RP3000022	3870	3871
	NT2RP3000171	C-NT2RP3000171	3872	3873
	NT2RP3000304	C-NT2RP3000304	3874	3875
	NT2RP3000378	C-NT2RP3000378	3876	3877
	NT2RP3000444	C-NT2RP3000444	3878	3879
40	NT2RP3000645	C-NT2RP3000645	3880	3881
	NT2RP3000676	C-NT2RP3000676	3882	3883
	NT2RP3000677	C-NT2RP3000677	3884	3885
	NT2RP3000789	C-NT2RP3000789	3886	3887
	NT2RP3000818	C-NT2RP3000818	3888	3889
45	NT2RP3000838	C-NT2RP3000838	3890	3891
	NT2RP3000921	C-NT2RP3000921	3892	3893
	NT2RP3001159	C-NT2RP3001159	3894	3895
	NT2RP3001271	C-NT2RP3001271	3896	3897
50	NT2RP3001542	C-NT2RP3001542	3898	3899
	NT2RP3001685	C-NT2RP3001685	3900	3901
	NT2RP3001976	C-NT2RP3001976	3902	3903
	NT2RP3002015	C-NT2RP3002015	3904	3905
	NT2RP3002281	C-NT2RP3002281	3906	3907
55	NT2RP3002286	C-NT2RP3002286	3908	3909
	NT2RP3002324	C-NT2RP3002324	3910	3911

EP 1 130 094 A2

	NT2RP3002353	C-NT2RP3002353	3912	3913
	NT2RP3002571	C-NT2RP3002571	3914	3915
	NT2RP3002664	C-NT2RP3002664	3916	3917
5	NT2RP3002737	C-NT2RP3002737	3918	3919
	NT2RP3002887	C-NT2RP3002887	3920	3921
	NT2RP3002900	C-NT2RP3002900	3922	3923
	NT2RP3002983	C-NT2RP3002983	3924	3925
	NT2RP3003473	C-NT2RP3003473	3926	3927
10	NT2RP3003532	C-NT2RP3003532	3928	3929
	NT2RP3004025	C-NT2RP3004025	3930	3931
	NT2RP3004067	C-NT2RP3004067	3932	3933
	NT2RP3004119	C-NT2RP3004119	3934	3935
15	NT2RP3004294	C-NT2RP3004294	3936	3937
	NT2RP3004345	C-NT2RP3004345	3938	3939
	NT2RP4000634	C-NT2RP4000634	3940	3941
	NT2RP4001001	C-NT2RP4001001	3942	3943
	NT2RP4001877	C-NT2RP4001877	3944	3945
20	NT2RP4001879	C-NT2RP4001879	3946	3947
	NT2RP4002187	C-NT2RP4002187	3948	3949
	NT2RP4002451	C-NT2RP4002451	3950	3951
	NT2RP4002750	C-NT2RP4002750	3952	3953
25	OVARC1000003	C-OVARC1000003	3954	3955
	OVARC1000313	C-OVARC1000313	3956	3957
	OVARC1000331	C-OVARC1000331	3958	3959
	OVARC1000553	C-OVARC1000553	3960	3961
	OVARC1000873	C-OVARC1000873	3962	3963
	OVARC1000995	C-OVARC1000995	3964	
30	OVARC1001260	C-OVARC1001260	3965	
	OVARC1001336	C-OVARC1001336	3966	3967
	OVARC1001570	C-OVARC1001570	3968	3969
	OVARC1001607	C-OVARC1001607	3970	3971
	OVARC1001833	C-OVARC1001833	3972	3973
35	OVARC1001952	C-OVARC1001952	3974	3975
	PLACE1000986	C-PLACE1000986	3976	
	PLACE1003407	C-PLACE1003407	3977	3978
	PLACE1004078	C-PLACE1004078	3979	3980
40	PLACE1004492	C-PLACE1004492	3981	3982
	PLACE1005539	C-PLACE1005539	3983	3984
	PLACE1005569	C-PLACE1005569	3985	3986
	PLACE1005601	C-PLACE1005601	3987	
	PLACE1005745	C-PLACE1005745	3988	3989
45	PLACE1005815	C-PLACE1005815	3990	3991
	PLACE1005927	C-PLACE1005927	3992	3993
	PLACE1006071	C-PLACE1006071	3994	3995
	PLACE1006073	C-PLACE1006073	3996	3997
	PLACE1006079	C-PLACE1006079	3998	3999
50	PLACE1006786	C-PLACE1006786	4000	
	PLACE1007077	C-PLACE1007077	4001	4002
	PLACE1007971	C-PLACE1007971	4003	
	PLACE1008282	C-PLACE1008282	4004	4005
	PLACE1008359	C-PLACE1008359	4006	4007
55	PLACE1008744	C-PLACE1008744	4008	4009

EP 1 130 094 A2

	PLACE1010445	C-PLACE1010445	4010	4011
	PLACE1010713	C-PLACE1010713	4012	4013
	nnnnnnnnnnnn	C-nnnnnnnnnnnn	4014	4015
5	PLACE1011181	C-PLACE1011181	4016	4017
	PLACE1011364	C-PLACE1011364	4018	4019
	PLACE3000181	C-PLACE3000181	4020	4021
	SKNMC1000014	C-SKNMC1000014	4022	4023
	SKNMC1000082	C-SKNMC1000082	4024	4025
10	THYRO1000964	C-THYRO1000964	4026	4027
	THYRO1001242	C-THYRO1001242	4028	4029
	THYRO1001608	C-THYRO1001608	4030	4031
	THYRO1001641	C-THYRO1001641	4032	4033
	THYRO1001770	C-THYRO1001770	4034	4035
15	Y79AA1000030	C-Y79AA1000030	4036	4037
	Y79AA1001212	C-Y79AA1001212	4038	4039
	Y79AA1001426	C-Y79AA1001426	4040	4041
	Y79AA1001427	C-Y79AA1001427	4042	4043
20	Y79AA1001523	C-Y79AA1001523	4044	4045
	Y79AA1001530	C-Y79AA1001530	4046	4047
	Y79AA1001592	C-Y79AA1001592	4048	4049
	Y79AA1001727	C-Y79AA1001727	4050	4051
	Y79AA1001787	C-Y79AA1001787	4052	4053
25	Y79AA1001793	C-Y79AA1001793	4054	4055
	Y79AA1001795	C-Y79AA1001795	4056	4057
	Y79AA1001799	C-Y79AA1001799	4058	4059
	Y79AA1001803	C-Y79AA1001803	4060	4061
	Y79AA1001863	C-Y79AA1001863	4062	4063
30	Y79AA1002058	C-Y79AA1002058	4064	4065
	Y79AA1002121	C-Y79AA1002121	4066	4067
	Y79AA1002213	C-Y79AA1002213	4068	4069
	Y79AA1002373	C-Y79AA1002373	4070	4071
	Y79AA1002376	C-Y79AA1002376	4072	4073
35	Y79AA1002378	C-Y79AA1002378	4074	4075
	Y79AA1002381	C-Y79AA1002381	4076	4077
	BNGH41000087	C-BNGH41000087	4078	4079
40	HEMBA1001886	C-HEMBA1001886	4080	4081
	HEMBA1004067	C-HEMBA1004067	4082	4083
	HEMBA1007226	C-HEMBA1007226	4084	4085
	HEMBB1000309	C-HEMBB1000309	4086	4087
	HEMBB1000567	C-HEMBB1000567	4088	4089
45	MAMMA1000102	C-MAMMA1000102	4090	4091
	MAMMA1001066	C-MAMMA1001066	4092	4093
	MAMMA1001094	C-MAMMA1001094	4094	4095
	MAMMA1001609	C-MAMMA1001609	4096	4097
	MAMMA1001901	C-MAMMA1001901	4098	
50	MAMMA1002091	C-MAMMA1002091	4099	4100
	NT2RM1000462	C-NT2RM1000462	4101	4102
	NT2RM1000542	C-NT2RM1000542	4103	4104
	NT2RM1000789	C-NT2RM1000789	4105	4106
	NT2RM1000855	C-NT2RM1000855	4107	4108
55	NT2RM1000899	C-NT2RM1000899	4109	4110

EP 1 130 094 A2

	NT2RP2000092	C-NT2RP2000092	4111	4112
	NT2RP2001538	C-NT2RP2001538	4113	4114
	NT2RP2001921	C-NT2RP2001921	4115	4116
5	NT2RP2003138	C-NT2RP2003138	4117	4118
	NT2RP2003302	C-NT2RP2003302	4119	4120
	NT2RP2003950	C-NT2RP2003950	4121	4122
	NT2RP2005535	C-NT2RP2005535	4123	4124
10	NT2RP2005774	C-NT2RP2005774	4125	4126
	NT2RP3000148	C-NT2RP3000148	4127	4128
	NT2RP3000232	C-NT2RP3000232	4129	4130
	NT2RP3000427	C-NT2RP3000427	4131	
15	NT2RP3000652	C-NT2RP3000652	4132	4133
	NT2RP3001650	C-NT2RP3001650	4134	4135
	NT2RP3002409	C-NT2RP3002409	4136	
	NT2RP3002411	C-NT2RP3002411	4137	4138
20	NT2RP3003448	C-NT2RP3003448	4139	
	NT2RP4002715	C-NT2RP4002715	4140	4141
	OVARC1000307	C-OVARC1000307	4142	4143
	PLACE1000907	C-PLACE1000907	4144	4145
25	PLACE1007081	C-PLACE1007081	4146	4147
	PLACE1010011	C-PLACE1010011	4148	4149
	PLACE3000213	C-PLACE3000213	4150	4151
	PLACE4000354	C-PLACE4000354	4152	4153
	PLACE4000455	C-PLACE4000455	4154	
30	THYRO1000776	C-THYRO1000776	4155	4156
	THYRO1001593	C-THYRO1001593	4157	4158
	Y79AA1000750	C-Y79AA1000750	4159	4160
	Y79AA1000888	C-Y79AA1000888	4161	4162
35	Y79AA1002129	C-Y79AA1002129	4163	4164
	Y79AA1002334	C-Y79AA1002334	4165	4166
	MAMMA1002224	C-MAMMA1002224	4167	
40	NT2RP1000271	C-NT2RP1000271	4168	4169
	NT2RP3000481	C-NT2RP3000481	4170	4171
	NT2RP3004481	C-NT2RP3004481	4172	4173
	HEMBA1006658	C-HEMBA1006658	4174	4175
	NT2RP2006099	C-NT2RP2006099	4176	4177
45	NT2RP2006580	C-NT2RP2006580	4178	4179

Homology search result 1

50 [0287] The result of the homology search in the SwissProt using the representative sequences of the 5'-ends.

Indicated are from the top,
the name of the representative sequence of the cluster,
55 definition of the top hit data,
the P-value: the length of the sequence used for comparison (nucleotide):similarity (%),
the organism of which the top hit data is obtained,
the Accession No. of the top hit data.

EP 1 130 094 A2

[0288] Homology search results of the representative sequences of the 5'-end cluster to the data in SwissProt database are shown only for the representative sequences of the cluster from which clones were selected based on the homology search results.

[0289] The P-value is the score which is determined by taking into account the statistic probability of occurrence between the two sequences, and generally low score reflects high similarity. (Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. (1990) "Basic local alignment search tool." J. Mol. Biol. 215:403-410; Gish, W. & States, D.J. (1993) "Identification of protein coding regions by database similarity search." Nature Genet. 3:266-272).

- 5
10
15
20
25
30
35
40
45
50
55
- HRIFA000016a
GLYCINE-RICH CELL WALL STRUCTURAL PROTEIN 1.8 PRECURSOR (GRP 1.8).
9.2e-05:178:32
PHASEOLUS VULGARIS (KIDNEY BEAN) (FRENCH BEAN).
P10496
- HRIFA000071a
CIRCUMSPOROZOITE PROTEB PRECURSOR (CS).
5.8e-05:194:29
PLASMODIUM SIMIUM.
Q03110
- HRIFA000116a
HYPOTHETICAL 68.7 KD PROTEIN ZK757.1 IN CHROMOSOME III.
6.2e-06:83:27
CAENORHABDITIS ELEGANS.
P34679
- HRIFA000123a
PATHOGENESIS-RELATED PROTEIN 1 PRECURSOR (PR-1).
6.2e-08:89:34
ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).
P33154
- HRIFA000264a
PROCOLLAGEN ALPHA 1(I) CHAIN PRECURSOR.
1.4e-06:231:34
GALLUS GALLUS (CHICKEN).
P02457
- HRIFA000327a
ATP-BINDING CASSETTE TRANSPORTER 1.
2.0e-16:238:31
MUS MUSCULUS (MOUSE).
P41233
- HRIFA000415a
PROLINE-RICH PROTEIN MP-2 PRECURSOR.
3.6e-06:120:35
MUS MUSCULUS (MOUSE).
P05142
- HRIFA000432a
PUTATIVE GENERAL NEGATIVE REGULATOR OF TRANSCRIPTION C16C9.04C.
2.2e-21:86:52
SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).
Q09818
- HRIFA000446a
HYPOTHETICAL 64.8 KD PROTEIN IN GDI1-COX15 INTERGENIC REGION.

EP 1 130 094 A2

- 2.5e-09:138:34
SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
P40085
- 5 HRIFA000553a
CARTILAGE MATRIX PROTEIN PRECURSOR (MATRILIN-1).
1.7e-27:117:48
GALLUS GALLUS (CHICKEN).
P05099
- 10 HRIFA000564a
ERYTHROCYTE BAND 7 INTEGRAL MEMBRANE PROTEIN (STOMATIN) (PROTEIN 7.2B).
2.9e-28:163:38
MUS MUSCULUS (MOUSE).
P54116
- 15 HRIFA000631a
ZINC FINGER PROTEIN 140.
8.2e-45:155:47
HOMO SAPIENS (HUMAN).
P52738
- 20 HRIFA000683a
FIBRILLIN 1 PRECURSOR.
4.8e-18:77:46
HOMO SAPIENS (HUMAN).
P35555
- 25 HRIFA000695a
*SALIVARY PROLINE-RICH PROTEIN PRECURSOR (CLONES CP3, CP4 AND CP5) [CONTAINS: BASIC PEP-
TIDE IB-6" PEPTIDE P-H].
4.0e-06:105:33
HOMO SAPIENS (HUMAN).
P04280
- 30 HRIFA000776a
FIBRILLIN 2 PRECURSOR.
1.6e-42:214:44
HOMO SAPIENS (HUMAN).
P35556
- 35 HRIFA000814a
ZINC FINGER PROTEIN 133.
4.4e-16:49:87
HOMO SAPIENS (HUMAN).
P52736
- 40 HRIFA000845a
PROCOLLAGEN ALPHA 1(I) CHAIN PRECURSOR.
6.0e-06:172:34
MUS MUSCULUS (MOUSE).
P11087
- 50 HRIFA001099a
CYCLIC-AMP-DEPENDENT TRANSCRIPTION FACTOR ATF-5 (FRAGMENT).
0.92:38:34
HOMO SAPIENS (HUMAN).
P18849
- 55

EP 1 130 094 A2

5 HRIFA001132a
 AGRIN PRECURSOR.
 1.3e-26:239:32
 GALLUS GALLUS (CHICKEN).
 P31696

10 HRIFA001138a
 CARTILAGE OLIGOMERIC MATRIX PROTEIN PRECURSOR (COMP).
 5.9e-114:147:83
 HOMO SAPIENS (HUMAN).
 P49747

15 HRIFA001200a
 TRANS-ACTING TRANSCRIPTIONAL PROTEIN ICP0 (P135 PROTEIN) (IER 2.9/ER2.6).
 0.035:119:34
 BOVINE HERPESVIRUS TYPE 1 (STRAIN JURA).
 P29128

20 HRIFA001337a
 LOW-DENSITY LIPOPROTEIN RECEPTOR PRECURSOR (LDL RECEPTOR).
 2.4e-17:98:42
 CRICETULUS GRISEUS (CHINESE HAMSTER).
 P35950

25 HRIFA001341a
 NEUROFILAMENT TRIPLET L PROTEIN (68 KD NEUROFILAMENT PROTEIN) (NF-L)
 (NF68).
 1.2e-102:248:87
 RATTUS NORVEGICUS (RAT).
 30 P19527

35 HRIFA001413a
 BACTENECIN 7 PRECURSOR (BAC7) (PR-59).
 0.0032:33:63
 BOS TAURUS (BOVINE).
 P19661

40 HRIFA001439a
 B-CELL GROWTH FACTOR PRECURSOR (BCGF-12 KD).
 0.00031:34:61
 HOMO SAPIENS (HUMAN).
 P20931

45 HRIFA001489a
 PROBABLE G PROTEIN-COUPLED RECEPTOR APJ.
 8.4e-65:105:72
 HOMO SAPIENS (HUMAN).
 P35414

50 HRIFA001558a
 HISTIDINE-RICH GLYCOPROTEIN PRECURSOR.
 0.0048:80:31
 PLASMODIUM LOPHURAE.
 P04929

55 HRIFA001712a
 PUTATIVE SERINE/THREONINE-PROTEIN KINASE PKWA (EC 2.7.1.-).
 2.5e-19:169:31

EP 1 130 094 A2

THERMOMONOSPORA CURVATA.
P49695

5 HRIFA001720a
ZINC FINGER PROTEIN 85 (ZINC FINGER PROTEIN HPF4) (HTF1).
1.4e-94:273:64
HOMO SAPIENS (HUMAN).
Q03923

10 HRIFA001866a
EARLY ANTIGEN PROTEIN D (EA-D).
0.10:93:34
EPSTEIN-BARR VIRUS (STRAIN B95-8) (HUMAN HERPESVIRUS 4).
P03191

15 HRIFA001942a
"PROCOLLAGEN-LYSINE,2-OXOGLUTARATE 5-DIOXYGENASE 1 PRECURSOR (EC 1.14.11.4) (LYSYL HY-
DROXYLASE 1) (LH1)."
4.7e-12:140:30
20 GALLUS GALLUS (CHICKEN).
P24802

HRIFA001971a
25 HYPOTHETICAL 46.3 KD PROTEIN IN PTA1-CDC24 INTERGENIC REGION.
2.5e-10:86:30
SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
P39727

HRIFA001972a
30 LAMININ ALPHA-1 CHAIN PRECURSOR (LAMININ A CHAIN).
0.10:100:34
MUS MUSCULUS (MOUSE).
P19137

35 HRIFA001975a
ACETYLCHOLINESTERASE PRECURSOR (EC 3.1.1.7).
6.5e-30:243:33
MUS MUSCULUS (MOUSE).
P21836

40 HRIFA001984a
"PROCOLLAGEN-LYSINE,2-OXOGLUTARATE 5-DIOXYGENASE 1 PRECURSOR (EC 1.14.11.4) (LYSYL HY-
DROXYLASE 1) (LH1)."
1.2e-11:140:30
45 GALLUS GALLUS (CHICKEN).
P24802

HRIFA002063a
50 GNS1 PROTEIN.
1.3e-05:127:30
SACCHAROMYCES CEREVISIAE (BAKERS YEAST).
P25358

HRIFA002102a
55 MUCIN 2 PRECURSOR (INTESTINAL MUCIN 2).
2.9e-07:241:30
HOMO SAPIENS (HUMAN).
Q02817

EP 1 130 094 A2

5 HRIFA002284a
 ACIDIC PROLINE-RICH PROTEIN PRECURSOR (CLONE PRP33).
 3.8e-05:104:34
 RATTUS NORVEGICUS (RAT).
 P04474

10 HRIFA002309a
 PUTATIVE SERINE/THREONINE-PROTEIN KINASE PKWA (EC 2.7.1.-).
 1.5e-08:110:37
 THERMOMONOSPORA CURVATA.
 P49695

15 HRIFA002384a
 GAP JUNCTION ALPHA-6 PROTEIN (CONNEXIN 45) (CX45).
 1.8e-31:94:42
 HOMO SAPIENS (HUMAN).
 P36383

20 HRIFA002503a
 N-ACETYLLACTOSAMINE SYNTHASE (EC 2.4.1.90) (N-ACETYLGALACTOSAMINE (BETA 1→4)GALACTOSYL-
 TRANSFERASE) (EC 2.4.1.38) (LACTOSE SYNTHASE A PROTEIN (EC 2.4.1.22)) (GALACTOSYLTRANS-
 FERASE) (GT).
 6.1e-92:246:67
 MUS MUSCULUS (MOUSE).
 25 P15535

30 HRIFA002689a
 TRANSCRIPTION FACTOR GATA-6 (GATA BINDING FACTOR-6) (DNA BINDING PROTEIN GATA-GT2).
 0.38:49:34
 RATTUS NORVEGICUS (RAT).
 P46153

35 HRIFA002694a
 ANTER-SPECIFIC PROLINE-RICH PROTEIN APG PRECURSOR.
 4.7e-05:93:37
 ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).
 P40602

40 HRIFA002743a
 BONE MORPHOGENETIC PROTEIN 1 PRECURSOR (EC 3.4.24.-) (BMP-1).
 1.2e-23:216:31
 HOMO SAPIENS (HUMAN).
 P13497

45 HRIFA002762a
 PROLINE-RICH PROTEIN MP-2 PRECURSOR.
 5.1e-09:129:41
 MUS MUSCULUS (MOUSE).
 P05142

50 HRIFA002766a
 FIBROMODULIN PRECURSOR (FM) (COLLAGEN-BINDING 59 KD PROTEIN).
 1.8e-12:139:34
 HOMO SAPIENS (HUMAN).
 55 Q06828

HRIFA002787a
 PROCOLLAGEN ALPHA 2(I) CHAIN PRECURSOR.

EP 1 130 094 A2

1.6e-10:124:37
HOMO SAPIENS (HUMAN).
P08123

5 HRIFA002805a
ZINC FINGER PROTEIN 140.
3.6e-23:43:74
HOMO SAPIENS (HUMAN).
P52738

10 HRIFA002891a
"FIBULIN-1, ISOFORM C PRECURSOR (BASEMENT-MEMBRANE PROTEIN 90) (BM-90)."
2.0e-41:239:39
MUS MUSCULUS (MOUSE).
15 Q08878

HRIFA002919a
BEM46 PROTEIN (FRAGMENT).
1.0e-12:171:32
20 SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).
P54069

HRIFA002980a
LOW-DENSITY LIPOPROTEIN RECEPTOR-RELATED PROTEIN 1 PRECURSOR (LRP) (ALPHA-2-MAC-
25 ROGLOBULIN RECEPTOR) (A2MR).
8.7e-32:202:37
GALLUS GALLUS (CHICKEN).
P98157

30 HRIFA003055a
PROLINE-RICH PROTEIN MP-2 PRECURSOR.
3.4e-08:175:29
MUS MUSCULUS (MOUSE).
P05142

35 HRIFA003063a
B-CELL LYMPHOMA 6 PROTEIN HOMOLOG.
2.8e-15:123:34
MUS MUSCULUS (MOUSE).
40 P41183

HRIFA003093a
PROLINE-RICH PROTEIN MP-2 PRECURSOR.
1.3e-11:142:37
45 MUS MUSCULUS (MOUSE).
P05142

HRIFA003340a
TRANS-ACTING TRANSCRIPTIONAL PROTEIN ICP0 (IMMEDIATE-EARLY PROTEIN IE110) (VMW110) (AL-
50 PHA-0 PROTEIN).
2.3e-05:200:31
HERPES SIMPLEX VIRUS (TYPE 1 / STRAIN 17).
P08393

55 HRIFA003357a
GLUCOSE REPRESSION MEDIATOR PROTEIN.
0.0023:190:26
SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).

EP 1 130 094 A2

P14922

- HRIFA003402a
COLLAGEN ALPHA 1(II) CHAIN (FRAGMENTS).
5 3.6e-05:194:27
BOS TAURUS (BOVINE).
P02459
- HRIFA003504a
10 CADHERIN-RELATED TUMOR SUPPRESSOR PRECURSOR (FAT PROTEIN).
1.4e-08:150:33
DROSOPHILA MELANOGASTER (FRUIT FLY).
P33450
- HRIFA003592a
15 CD9 ANTIGEN.
0.0053:24:70
BOS TAURUS (BOVINE).
P30932
- HRIFA003635a
20 "MANNOSYL-OLIGOSACCHARIDE ALPHA-1,2-MANNOSIDASE ISOFORM 1 (EC 3.2.1.113) (MAN(9)-ALPHA-MANNOSIDASE)."
5.3e-45:239:43
25 DROSOPHILA MELANOGASTER (FRUIT FLY).
P53624
- HRIFA003640a
PROCYCLIC FORM SPECIFIC POLYPEPTIDE A-BETA PRECURSOR (PROCYCLIN) (PARP A-BETA).
30 0.00018:28:64
TRYPANOSOMA BRUCEI BRUCEI.
P09791
- HRIFA003883a
35 TRANSCRIPTIONAL REPRESSOR PROTEIN YY1 (YIN AND YANG 1) (YY-1) (DELTA TRANSCRIPTION FACTOR) (NF-E1) (UCR-MOTIF DNA-BINDING PROTEIN).
1.0:57:35
MUS MUSCULUS (MOUSE).
Q00899
- HRIFA003892a
40 MULTIDRUG RESISTANCE PROTEIN 2 (MULTIDRUG-EFFLUX TRANSPORTER 2).
6.5e-08:144:25
BACILLUS SUBTILIS.
45 P39843
- HRIFA003946a
PROLINE-RICH PROTEIN MP-2 PRECURSOR.
1.4e-06:85:37
50 MUS MUSCULUS (MOUSE).
P05142
- HRIFA004006a
55 ZINC FINGER PROTEIN 140.
6.2e-20:83:66
HOMO SAPIENS (HUMAN).
P52738

EP 1 130 094 A2

- 5 HRIFA004034a
B-CELL LYMPHOMA 3-ENCODED PROTEIN (BCL-3 PROTEIN).
1.4e-15:192:32
HOMO SAPIENS (HUMAN).
P20749
- 10 HRIFA004112a
CADHERIN-RELATED TUMOR SUPPRESSOR PRECURSOR (FAT PROTEIN).
7.2e-26:193:37
DROSOPHILA MELANOGASTER (FRUIT FLY).
P33450
- 15 HRIFA004162a
ERYTHROCYTE BAND 7 INTEGRAL MEMBRANE PROTEIN (STOMATIN) (PROTEIN 7.2B).
3.6e-10:117:29
MUS MUSCULUS (MOUSE).
P54116
- 20 HRIFA004401a
LACTOSE OPERON REPRESSOR.
1.1e-07:36:86
ESCHERICHIA COLI.
P03023
- 25 HRIFA004426a
ATRIAL GLAND-SPECIFIC ANTIGEN PRECURSOR (AGSA).
5.1e-11:85:41
APLYSIA CALIFORNICA (CALIFORNIA SEA HARE).
P15287
- 30 HRIFA004490a
BETA-GALACTOSIDASE PRECURSOR (EC 3.2.1.23) (LACTASE) (ACID BETA-GALACTOSIDASE).
5.3e-19:101:44
MUS MUSCULUS (MOUSE).
35 P23780
- 40 HRIFA004523a
GLYCOSYLTRANSFERASE ALG2 (EC 2.4.1.-).
2.6e-36:180:43
SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
P43636
- 45 HRIFA004663a
T-CELL-SPECIFIC TRANSCRIPTION FACTOR 1 (TCF-1) (T-CELL FACTOR 1) (TRANSCRIPTION FACTOR-7).
1.2e-40:112:75
MUS MUSCULUS (MOUSE).
Q00417
- 50 HRIFA004696a
PROTEIN TRANSPORT PROTEIN SEC61 ALPHA SUBUNIT.
1.1e-62:145:84
CANIS FAMILIARIS (DOG).
P38377
- 55 HRIFA004714a
HYPOTHETICAL 36.7 KD PROTEIN AH6.2 IN CHROMOSOME II.
2.3e-50:127:54
CAENORHABDITIS ELEGANS.

EP 1 130 094 A2

Q09201

- HRIFA004745a
MITOCHONDRIAL RNA SPLICING PROTEIN MSR4.
5 5.0e-17:107:43
SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
P23500
- HRIFA004780a
10 EXTENSIN PRECURSOR (PROLINE-RICH GLYCOPROTEIN).
7.2e-07:142:30
ZEA MAYS (MAIZE).
P14918
- HRIFA004919a
15 GLYCINE-RICH CELL WALL STRUCTURAL PROTEIN 1.8 PRECURSOR (GRP 1.8).
1.5e-25:156:46
PHASEOLUS VULGARIS (KIDNEY BEAN) (FRENCH BEAN).
P10496
- HRIFA005072a
20 GLYCINE-RICH CELL WALL STRUCTURAL PROTEIN (CLONE W10-1) (FRAGMENT).
8.3e-05:24:62
LYCOPERSICON ESCULENTUM (TOMATO).
25 Q01157
- HRIFA005102a
A-AGGLUTININ ATTACHMENT SUBUNIT PRECURSOR.
2.5e-07:188:31
30 SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
P32323
- HRIFA005184a
35 CYTOCHROME B5.
3.4e-11:117:29
SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
P40312
- HRIFA005214a
40 TRANS-ACTING TRANSCRIPTIONAL PROTEIN ICP0 (VMW118 PROTEIN).
5.9e-05:141:33
HERPES SIMPLEX VIRUS (TYPE 2 / STRAIN HG52).
P28284
- HRIFA005231a
45 ORM1 PROTEIN.
1.7e-18:137:35
SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
P53224
- HRIFA005240a
50 ZINC FINGER PROTEIN 85 (ZINC FINGER PROTEIN HPF4) (HTF1).
6.3e-81:194:70
HOMO SAPIENS (HUMAN).
55 Q03923
- HRIFA005255a
HYPOTHETICAL 57.1 KD PROTEIN IN MAP2-TEL1 INTERGENIC REGION.

EP 1 130 094 A2

- 1.5e-07:202:24
SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
P38176
- 5 HRIFA005271a
MITOCHONDRIAL PHOSPHATE CARRIER PROTEIN PRECURSOR.
1.2e-55:86:81
HOMO SAPIENS (HUMAN).
Q00325
- 10 HRIFA005296a
INSULIN PROMOTER FACTOR 1 (IPF-1) (ISLET/DUODENUM HOMEODOMAIN-1) (IDX-1) (SOMATOSTATIN
TRANSACTIVATING FACTOR-1) (STF-1) (PANCREAS/DUODENUM HOMEODOMAIN-1) (GLUCOSE SENSITIVE
FACTOR) (GSF).
15 0.82:90:34
HOMO SAPIENS (HUMAN).
P52945
- 20 HRIFA005300a
SKIN SECRETORY PROTEIN XP2 PRECURSOR (APEG PROTEIN).
1.6e-07:178:30
XENOPUS LAEVIS (AFRICAN CLAWED FROG).
P17437
- 25 HRIFA005369a
EBNA-1 NUCLEAR PROTEIN.
2.3e-07:101:39
EPSTEIN-BARR VIRUS (STRAIN B95-8) (HUMAN HERPESVIRUS 4).
P03211
- 30 HRIFA005372a
CCAAT-BINDING TRANSCRIPTION FACTOR SUBUNIT A (CBF-A) (NF-Y PROTEIN CHAIN B) (NF-YB) (CAAT-
BOX DNA BINDING PROTEIN SUBUNIT B).
1.1e-14:97:38
35 PETROMYZON MARINUS (SEA LAMPREY).
P25210
- 40 HRIFA005392a
SYNDECAN-2 PRECURSOR (FIBROGLYCAN) (HEPARAN SULFATE PROTEOGLYCAN CORE PROTEIN)
(HSPG) (SYND2).
1.3e-50:126:84
HOMO SAPIENS (HUMAN).
P34741
- 45 HRIFA005409a
HIGH-AFFINITY CATIONIC AMINO ACID TRANSPORTER-1 (CAT-1) (CAT1) (SYSTEM Y+ BASIC AMINO ACID
TRANSPORTER) (ECOTROPIC RETROVIRAL LEUKEMIA RECEPTOR HOMOLOG) (ERR) (ECOTROPIC RET-
ROVIRUS RECEPTOR HOMOLOG).
7.1e-66:197:64
50 HOMO SAPIENS (HUMAN).
P30825
- 55 HRIFA005420a
INTERFERON-RELATED PROTEIN PC4 (TPA INDUCED SEQUENCE 7) (TIS7 PROTEIN).
1.5e-33:221:41
MUS MUSCULUS (MOUSE).
P19182

EP 1 130 094 A2

HRIFA005438a
 SUCCINATE DEHYDROGENASE [UBIQUINONE] FLAVOPROTEIN SUBUNIT PRECURSOR (EC 1.3.5.1) (FP)
 (FLAVOPROTEIN SUBUNIT OF COMPLEX II).
 6.4e-71:175:68
 5 HOMO SAPIENS (HUMAN).
 P31040

HRIFA005462a
 CARBONIC ANHYDRASE VI (EC 4.2.1.1) (SALIVARY) (CARBONATE DEHYDRATASE VI).
 10 1.4e-19:137:37
 OVIS ARIES (SHEEP).
 P08060

HRIFA005500a
 15 EBNA-1 NUCLEAR PROTEIN.
 0.00042:54:50
 EPSTEIN-BARR VIRUS (STRAIN B95-8) (HUMAN HERPESVIRUS 4).
 P03211

HRIFA005540a
 20 CYCLIC-AMP-DEPENDENT TRANSCRIPTION FACTOR ATF-6 (FRAGMENT).
 0.12:47:29
 HOMO SAPIENS (HUMAN).
 P18850

HRIFA005644a
 25 VACUOLAR ATP SYNTHASE SUBUNIT AC45 PRECURSOR (EC 3.6.1.34) (V-ATPASE AC45 SUBUNIT).
 1.2e-102:233:87
 BOS TAURUS (BOVINE).
 30 P40682

HRIFA005702a
 CELL SURFACE GLYCOPROTEIN MUC18 PRECURSOR (MELANOMA-ASSOCIATED ANTIGEN MUC18)
 (MELANOMA-ASSOCIATED ANTIGEN A32) (S-ENDO 1 ENDOTHELIAL- ASSOCIATED ANTIGEN) (CD146 AN-
 35 TIGEN) (MELANOMA ADHESION MOLECULE).
 8.7e-05:174:28
 HOMO SAPIENS (HUMAN).
 P43121

HRIFA005720a
 40 F-SPONDIN PRECURSOR.
 8.9e-12:155:31
 XENOPUS LAEVIS (AFRICAN CLAWED FROG).
 P35447

HRIFA005728a
 45 SPORULATION-SPECIFIC PROTEIN 1 (EC 2.7.1.-).
 1.7e-05:126:29
 SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
 50 P08458

HRIFA005732a
 PUTATIVE SERINE/THREONINE-PROTEIN KINASE PKWA (EC 2.7.1.-).
 4.4e-26:159:38
 55 THERMOMONOSPORA CURVATA.
 P49695

HRIFA005760a

EP 1 130 094 A2

- FUCOSYLGLYCOPROTEIN ALPHA-N-ACETYL GALACTOSAMINYLTRANSFERASE (EC 2.4.1.40) (HISTO-BLOOD GROUP A TRANSFERASE) (A TRANSFERASE) / FUCOSYLGLYCOPROTEIN 3-ALPHA-GALACTOSYLTRANSFERASE (EC 2.4.1.37) (HISTO-BLOOD GROUP B TRANSFERASE) (B TRANSFERASE) (NAGAT).
3.8e-15:53:54
5 HOMO SAPIENS (HUMAN).
P16442
- HRIFA005781a
10 ESTRADIOL 17 BETA-DEHYDROGENASE 3 (EC 1.1.1.62) (17-BETA-HSD 3) (TESTICULAR 17-BETA-HYDROXYSTEROID DEHYDROGENASE).
5.2e-47:228:47
HOMO SAPIENS (HUMAN).
P37058
- 15 HRIFA005944a
PROCOLLAGEN ALPHA 1(II) CHAIN PRECURSOR [CONTAINS: CHONDROCALCIN].
2.5e-06:142:35
MUS MUSCULUS (MOUSE).
P28481
- 20 HRIFA006183a
ZINC FINGER PROTEIN 136.
1.3e-42:129:62
HOMO SAPIENS (HUMAN).
25 P52737
- HRIFA006250a
HOMEOTIC GENE REGULATOR (BRAHMA PROTEIN).
0.0038:75:37
30 DROSOPHILA MELANOGASTER (FRUIT FLY).
P25439
- HRIFA006298a
EBNA-1 NUCLEAR PROTEIN.
35 1.4e-05:80:42
EPSTEIN-BARR VIRUS (STRAIN B95-8) (HUMAN HERPESVIRUS 4).
P03211
- HRIFA006448a
40 SKIN SECRETORY PROTEIN XP2 PRECURSOR (APEG PROTEIN).
8.5e-05:183:28
XENOPUS LAEVIS (AFRICAN CLAWED FROG).
P17437
- 45 HRIFA006494a
AXONIN-1 PRECURSOR (AXONAL GLYCOPROTEIN TAG-1) (TRANSIENT AXONAL GLYCOPROTEIN 1).
1.2e-18:201:33
HOMO SAPIENS (HUMAN).
Q02246
- 50 HRIFA006510a
CORNICHON PROTEIN.
6.0e-53:144:66
DROSOPHILA MELANOGASTER (FRUIT FLY).
55 P49858
- HRIFA006566a
CCAAT-BINDING TRANSCRIPTION FACTOR SUBUNIT A (CBF-A) (NF-Y PROTEIN CHAIN B) (NF-YB) (CAAT-

EP 1 130 094 A2

BOX DNA BINDING PROTEIN SUBUNIT B).
6.6e-15:97:38
PETROMYZON MARINUS (SEA LAMPREY).
P25210

5

HRIFA006572a
PROCOLLAGEN ALPHA 1(I) CHAIN PRECURSOR.
7.2e-05:158:29
MUS MUSCULUS (MOUSE).
10 P11087

HRIFA006586a
HYPOTHETICAL 53.3 KD PROTEIN IN HXT8-CAN1 INTERGENIC REGION.
1.3e-13:219:26
15 SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
P39981

HRIFA006596a
POTENTIAL CAAX PRENYL PROTEASE 1 (EC 3.4.24.-) (PRENYL PROTEIN- SPECIFIC ENDOPROTEASE 1)
20 (PPSEP 1).
7.2e-22:241:32
SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).
Q10071

25 HRIFA006609a
PANCREATIC HORMONE PRECURSOR (PANCREATIC POLYPEPTIDE) (PP).
0.61:28:46
"GALLUS GALLUS (CHICKEN), AND MELEAGRIS GALLOPAVO (COMMON TURKEY)."
P01306

30 HRIFA006633a
COLLAGEN ALPHA 1(XVI) CHAIN PRECURSOR.
7.8e-07:170:34
HOMO SAPIENS (HUMAN).
35 Q07092

HRIFA006642a
AMALGAM PROTEIN PRECURSOR.
1.5e-09:185:28
40 DROSOPHILA MELANOGASTER (FRUIT FLY).
P15364

HRIFA006649a
ZINC FINGER PROTEIN 85 (ZINC FINGER PROTEIN HPF4) (HTF1).
45 1.7e-50:166:50
HOMO SAPIENS (HUMAN).
Q03923

HRIFA006667a
50 ZINC FINGER PROTEIN 85 (ZINC FINGER PROTEIN HPF4) (HTF1).
6.8e-45:180:43
HOMO SAPIENS (HUMAN).
Q03923

55 HRIFA006730a
SYG1 PROTEIN.
1.8e-14:164:35
SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).

EP 1 130 094 A2

P40528

HRIFA006798a
SKIN SECRETORY PROTEIN XP2 PRECURSOR (APEG PROTEIN).
0.22:149:34
XENOPUS LAEVIS (AFRICAN CLAWED FROG).
P17437

5

HRIFA006926a
SYNAPTOTAGMIN IV.
3.6e-19:168:38
RATTUS NORVEGICUS (RAT).
P50232

10

HRIFA007013a
MIC1 PROTEIN.
1.4e-13:115:38
SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
P53258

15

20

HRIFA007032a
CCAAT DISPLACEMENT PROTEIN (HOMEBOX PROTEIN CLOX) (CLOX-1) (FRAGMENT).
0.00013:92:35
CANIS FAMILIARIS (DOG).
P39881

25

HRIFA007068a
EBNA-1 NUCLEAR PROTEIN.
7.0e-10:145:33
EPSTEIN-BARR VIRUS (STRAIN B95-8) (HUMAN HERPESVIRUS 4).
P03211

30

HRIFA007152a
TRANSCRIPTION FACTOR SOX-4.
0.90:47:44
HOMO SAPIENS (HUMAN).
Q06945

35

HRIFA007219a
THROMBOSPONDIN 3 PRECURSOR.
1.3e-105:209:88
HOMO SAPIENS (HUMAN).
P49746

40

HRIFA007228a
HYPOTHETICAL 53.3 KD PROTEIN IN HXT8-CAN1 INTERGENIC REGION.
2.3e-11:174:24
SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
P39981

45

50

HRIFA007243a
PROBABLE CALCIUM-TRANSPORTING ATPASE 6 (EC 3.6.1.38).
3.0e-18:163:36
SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
P39986

55

HRIFA007244a
EXTENSIN PRECURSOR (CELL WALL HYDROXYPROLINE-RICH GLYCOPROTEIN).

EP 1 130 094 A2

- 4.2e-05:81:33
NICOTIANA TABACUM (COMMON TOBACCO).
P13983
- 5 HRIFA007256a
DEATH-ASSOCIATED PROTEIN KINASE 1 (EC 2.7.1.-) (DAP KINASE 1).
2.3e-77:186:75
HOMO SAPIENS (HUMAN).
P53355
- 10 HRIFA007262a
PAIRED AMPHIPATHIC HELIX PROTEIN.
1.3e-06:152:26
SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
15 P22579
- HRIFA007352a
5'-TG-3'INTERACTING FACTOR (HOMEBOX PROTEIN TGIF).
4.2e-36:146:57
20 HOMO SAPIENS (HUMAN).
Q15583
- HRIFA007424a
F-SPONDIN PRECURSOR.
25 8.9e-34:84:89
RATTUS NORVEGICUS (RAT).
P35446
- HRIFA007435a
30 PROTEIN KINASE CEK1 (EC 2.7.1.-).
1.0e-37:159:53
SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).
P38938
- 35 HRIFA007463a
HST1 PROTEIN (HOMOLOGOUS TO SIR2 PROTEIN 1).
4.8e-32:85:48
SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
P53685
- 40 HRIFA007493a
UBIQUITIN-CONJUGATING ENZYME E2-28.4 KD (EC 6.3.2.19) (UBIQUITIN- PROTEIN LIGASE) (UBIQUITIN
CARRIER PROTEIN).
1.2e-47:171:56
45 SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
P33296
- HRIFA007512a
50 EXTENSIN PRECURSOR (CELL WALL HYDROXYPROLINE-RICH GLYCOPROTEIN).
8.0e-07:173:28
NICOTIANA TABACUM (COMMON TOBACCO).
P13983
- 55 HRIFA007532a
"CALPAIN P94, LARGE [CATALYTIC] SUBUNIT (EC 3.4.22.17) (CALCIUM-ACTIVATED NEUTRAL PROTEIN-
ASE) (CANP) (P94 PROTEIN) (MUSCLE-SPECIFIC CALCIUM-ACTIVATED NEUTRAL PROTEASE 3 LARGE
SUBUNIT)."
1.8e-10:110:37

EP 1 130 094 A2

HOMO SAPIENS (HUMAN).
P20807

5 HRIFA007547a
TRANS-ACTING TRANSCRIPTIONAL PROTEIN ICP0 (P135 PROTEIN) (IER 2.9/ER2.6).
0.068:51:45
BOVINE HERPESVIRUS TYPE 1 (STRAIN K22).
P29836

10 HRIFA007565a
COLLAGEN ALPHA 1(X) CHAIN PRECURSOR.
5.1e-08:121:37
HOMO SAPIENS (HUMAN).
Q03692

15 HRIFA007571a
ORM1 PROTEIN.
5.8e-17:106:36
SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
20 P53224

HRIFA007659a
HYPOTHETICAL 51.5 KD PROTEIN D2024.3 IN CHROMOSOME IV.
2.5e-47:213:41
25 CAENORHABDITIS ELEGANS.
P49191

HRIFA007722a
HYPOTHETICAL 24.5 KD PROTEIN IN SAP185-BCK1 INTERGENIC REGION.
7.7e-13:146:32
30 SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
P40857

HRIFA007728a
35 EXTENSIN PRECURSOR (CELL WALL HYDROXYPROLINE-RICH GLYCOPROTEIN).
9.1e-05:124:31
NICOTIANA TABACUM (COMMON TOBACCO).
P13983

40 HRIFA007745a
ACETYLCHOLINESTERASE PRECURSOR (EC 3.1.1.7) (ACHE).
7.0e-15:109:36
TORPEDO CALIFORNICA (PACIFIC ELECTRIC RAY).
P04058

45 HRIFA007829a
GLYCINE-RICH CELL WALL STRUCTURAL PROTEIN (CLONE W10-1) (FRAGMENT).
0.00045:16:68
LYCOPERSICON ESCULENTUM (TOMATO).
50 Q01157

HRIFA007909a
COLLAGEN ALPHA 1(I) CHAIN (FRAGMENTS).
6.1e-06:173:34
55 BOS TAURUS (BOVINE).
P02453

HRIFA007985a

EP 1 130 094 A2

T-CELL RECEPTOR BETA CHAIN PRECURSOR (ANA 11).
0.00079:97:37
ORYCTOLAGUS CUNICULUS (RABBIT).
P06333

5

HRIFA008000a
"DIHYDROPYRIDINE-SENSITIVE L-TYPE, CALCIUM CHANNEL ALPHA-2/DELTA SUBUNITS PRECURSOR."
1.6e-37:165:42
ORYCTOLAGUS CUNICULUS (RABBIT).
P13806

10

HRIFA008174a
COLLAGEN 1(X) CHAIN PRECURSOR.
4.5e-05:215:28
BOS TAURUS (BOVINE).
P23206

15

HRIFA008186a
ESTRADIOL 17 BETA-DEHYDROGENASE 3 (EC 1.1.1.62) (17-BETA-HSD 3) (TESTICULAR 17-BETA-HY-
DROXYSTEROID DEHYDROGENASE).
2.1e-25:118:46
HOMO SAPIENS (HUMAN).
P37058

20

HRIFA008200a
ENDOSOMAL P24A PROTEIN PRECURSOR (70 KD ENDOMEMBRANE PROTEIN) (PHEROMONE ALPHA-
FACTOR TRANSPORTER) (ACIDIC 24 KD LATE ENDOCYTIC INTERMEDIATE COMPONENT).
7.9e-17:139:36
SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
P32802

25

HRIFA008212a
A-AGGLUTININ ATTACHMENT SUBUNIT PRECURSOR.
0.035:135:28
SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
P32323

30

HRIFA008252a
PROLINE-RICH PROTEIN MP-2 PRECURSOR.
0.00015:128:32
MUS MUSCULUS (MOUSE).
P05142

35

HRIFA008284a
NEURAL CELL ADHESION MOLECULE L1 PRECURSOR (N-CAM L1).
3.9e-18:153:30
HOMO SAPIENS (HUMAN).
P32004

40

HRIFA008314a
HYPOTHETICAL 149.7 KD PROTEIN IN IRE1-KSP1 INTERGENIC REGION.
2.1e-18:99:47
SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
P38800

45

HRIFA008362a
PHOSPHATIDYLCHOLINE-STEROL ACYLTRANSFERASE PRECURSOR (EC 2.3.1.43) (LECITHIN-CHOLE-
STEROL ACYLTRANSFERASE) (PHOSPHOLIPID-CHOLESTEROL ACYLTRANSFERASE) (FRAGMENT).

50

55

EP 1 130 094 A2

9.1e-42:135:57
GALLUS GALLUS (CHICKEN).
P53760

5 HRIFA008426a
HOMEODOMAIN PROTEIN AKR (AVIAN KNOTTED-RELATED PROTEIN).
1.3e-08:104:45
GALLUS GALLUS (CHICKEN).
Q90655

10 HRIFA008459a
CARBON CATABOLITE DEREPRESSING PROTEIN KINASE (EC 2.7.1.-).
5.5e-15:96:40
SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
15 P06782

HRIFA008483a
PROBABLE LONG-CHAIN FATTY ACID TRANSPORT PROTEIN.
7.4e-26:154:41

20 SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
P38225

HRIFA008547a
ZINC FINGER PROTEIN 136.
7.2e-57:228:50
HOMO SAPIENS (HUMAN).
25 P52737

HRIFA008596a
30 SKIN SECRETORY PROTEIN XP2 PRECURSOR (APEG PROTEIN).
1.6e-05:97:35
XENOPUS LAEVIS (AFRICAN CLAWED FROG).
P17437

35 HRIFA008611a
NPL1 PROTEIN (SEC63 PROTEIN).
8.1e-15:113:38
SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
P14906

40 HRIFA008661a
GALACTOSE-PROTON SYMPORT (GALACTOSE TRANSPORTER).
2.7e-16:184:29
ESCHERICHIA COLI.

45 P37021

HRIFA008717a
SERINE/THREONINE-PROTEIN KINASE NRK1 (EC 2.7.1.-) (N-RICH KINASE 1).
6.9e-32:198:41

50 SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
P38692

HRIFA008784a
55 HYPOTHETICAL 26.3 KD PROTEIN IN OYE2-GND1 INTERGENIC REGION.
2.2e-16:93:47
SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
P38869

EP 1 130 094 A2

HRIFA008790a
HYPOTHETICAL 15.7 KD PROTEIN IN NUP85-SSC1 INTERGENIC REGION.
4.2e-08:121:32
SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
5 P47111

HRIFA008976a
ACROSIN PRECURSOR (EC 3.4.21.10).
0.31:20:70
10 HOMO SAPIENS (HUMAN).
P10323

HRIFA008981a
ZINC FINGER PROTEIN 85 (ZINC FINGER PROTEIN HPF4) (HTF1).
15 1.0e-84:126:74
HOMO SAPIENS (HUMAN).
Q03923

HRIFA008989a
20 TRANS-ACTING TRANSCRIPTIONAL PROTEIN ICP0 (IMMEDIATE-EARLY PROTEIN IE110) (VMW110) (AL-
PHA-0 PROTEIN).
1.2e-05:134:33
HERPES SIMPLEX VIRUS (TYPE 1 / STRAIN 17).
P08393
25

HRIFA009071 a
CELLULAR TUMOR ANTIGEN P53 (PHOSPHOPROTEIN P53).
0.14:104:31
HOMO SAPIENS (HUMAN).
30 P04637

HRIFA009101a
ZINC FINGER PROTEIN 136.
6.5e-47:126:67
35 HOMO SAPIENS (HUMAN).
P52737

HRIFA009123a
SKIN SECRETORY PROTEIN XP2 PRECURSOR (APEG PROTEIN).
40 0.010:127:35
XENOPUS LAEVIS (AFRICAN CLAWED FROG).
P17437

HRIFA009136a
45 REGULATORY PROTEIN E2.
0.032:100:37
HUMAN PAPILLOMAVIRUS TYPE 25.
P36787

HRIFA009171a
50 BUTYROPHILIN PRECURSOR (BT).
1.6e-15:168:31
BOS TAURUS (BOVINE).
P18892
55

HRIFA009220a
HYPOTHETICAL 43.7 KD PROTEIN C24B11.08C IN CHROMOSOME I.
2.2e-48:268:41

EP 1 130 094 A2

SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).
Q09895

5 HRIFA009339a
PROCOLLAGEN ALPHA 2(I) CHAIN PRECURSOR.
0.63:57:35
MUS MUSCULUS (MOUSE).
Q01149

10 HRIFA009451a
METALLOPROTEINASE INHIBITOR 1 PRECURSOR (TIMP-1) (ERYTHROID POTENTIATING ACTIVITY) (EPA)
(TISSUE INHIBITOR OF METALLOPROTEINASES) (FIBROBLAST COLLAGENASE INHIBITOR) (COLLA-
GENASE INHIBITOR).
1.7e-57:163:73
15 HOMO SAPIENS (HUMAN).
P01033

HRIFA009482a
20 BETA-GALACTOSIDASE PRECURSOR (EC 3.2.1.23) (LACTASE) (ACID BETA-GALACTOSIDASE).
7.7e-25:86:59
MUS MUSCULUS (MOUSE).
P23780

HRIFA009578a
25 HYPOTHETICAL 24.5 KD PROTEIN IN SAP185-BCK1 INTERGENIC REGION.
8.8e-10:199:26
SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
P40857

30 HRIFA009762a
CCAAT DISPLACEMENT PROTEIN (CDP) (CDP2) (FRAGMENT).
0.17:116:32
RATTUS NORVEGICUS (RAT).
P53565

35 HRIFA009783a
HYPOTHETICAL 85.7 KD PROTEIN C13G6.03 IN CHROMOSOME I.
6.2e-48:231:48
SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).
40 Q09782

HRIFA009825a
ANTER-SPECIFIC PROLINE-RICH PROTEIN APG PRECURSOR.
4.0e-06:70:38
45 ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).
P40602

HRIFA009852a
50 "NEURAL CELL ADHESION MOLECULE 1, LARGE ISOFORM PRECURSOR (N-CAM 180) [CONTAINS: N-CAM
140]."
4.0e-07:198:27
XENOPUS LAEVIS (AFRICAN CLAWED FROG).
P16170

55 HRIFA009881a
EXTENSIN PRECURSOR (PROLINE-RICH GLYCOPROTEIN).
1.5e-11:106:35
SORGHUM VULGARE (SORGHUM).

EP 1 130 094 A2

P24152

HRIFA009983a
G-BOX BINDING FACTOR (GBF).
5 3.8e-10:156:30
DICTYOSTELIUM DISCOIDEUM (SLIME MOLD).
P36417

HRIFA010005a
10 "M PROTEIN, SEROTYPE 49 PRECURSOR."
1.6e-05:183:27
STREPTOCOCCUS PYOGENES.
P16947

HRIFA010078a
15 HYPOTHETICAL 57.5 KD PROTEIN IN VMA7-RPS25A INTERGENIC REGION.
4.7e-05:194:31
SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
P53214

HRIFA010085a
20 ZINC FINGER PROTEIN 85 (ZINC FINGER PROTEIN HPF4) (HTF1).
2.9e-92:243:69
HOMO SAPIENS (HUMAN).
25 Q03923

HRIFA010090a
N-ACETYLGLUCOSAMINE-6-SULFATASE PRECURSOR (EC 3.1.6.14) (G6S) (GLUCOSAMINE-6-SULFA-
TASE).
30 6.7e-16:78:51
HOMO SAPIENS (HUMAN).
P15586

HRIFA010130a
35 DOLICHYL-PHOSPHATE-MANNOSE--PROTEIN MANNOSYLTRANSFERASE 4 (EC 2.4.1.109).
5.6e-13:99:34
SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
P46971

HRIFA010152a
40 "ADENYLATE CYCLASE, TYPE V (EC 4.6.1.1) (ATP PYROPHOSPHATE-LYASE) (CA(2+)-INHIBITABLE ADE-
NYLYL CYCLASE)."
2.3e-05:73:43
CANIS FAMILIARIS (DOG).
45 P30803

HRIFA010176a
HEPATOCTE NUCLEAR FACTOR 3-BETA (HNF-3B).
0.066:105:31
50 MUS MUSCULUS (MOUSE).
P35583

HRIFA010301a
55 ANTER-SPECIFIC PROLINE-RICH PROTEIN APG PRECURSOR.
1.1e-09:120:34
ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).
P40602

EP 1 130 094 A2

- HRIFA010319a
DOPAMINE-BETA-MONOOXYGENASE PRECURSOR (EC 1.14.17.1) (DOPAMINE BETA- HYDROXYLASE)
(DBH).
4.8e-23:185:32
5 RATTUS NORVEGICUS (RAT).
Q05754
- HRIFA010361a
10 SKIN SECRETORY PROTEIN XP2 PRECURSOR (APEG PROTEIN).
2.6e-08:136:32
XENOPUS LAEVIS (AFRICAN CLAWED FROG).
P17437
- HRIFA010394a
15 HYPOTHETICAL 51.5 KD PROTEIN D2024.3 IN CHROMOSOME IV.
3.3e-36:144:47
CAENORHABDITIS ELEGANS.
P49191
- HRIFA010425a
20 A-AGGLUTININ ATTACHMENT SUBUNIT PRECURSOR.
1.9e-09:199:29
SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
P32323
- HRIFA010460a
25 TRANSCRIPTIONAL ACTIVATOR FE65.
2.3e-27:101:54
RATTUS NORVEGICUS (RAT).
30 P46933
- HRIFA010466a
SKIN SECRETORY PROTEIN XP2 PRECURSOR (APEG PROTEIN).
5.3e-07:123:34
35 XENOPUS LAEVIS (AFRICAN CLAWED FROG).
P17437
- HRIFA010490a
40 TRANS-ACTING TRANSCRIPTIONAL PROTEIN ICP0 (EARLY PROTEIN 0) (EP0).
0.0031:118:30
PSEUDORABIES VIRUS (STRAIN INDIANA-FUNKHAUSER / BECKER) (PRV).
P29129
- HRIFA010736a
45 PROTEIN Q300.
0.018:14:85
MUS MUSCULUS (MOUSE).
Q02722
- HRIFA010790a
50 RENAL SODIUM-DEPENDENT PHOSPHATE TRANSPORT PROTEIN 2 (SODIUM/PHOSPHATE COTRANS-
PORTER 2) (NA(+)/PI COTRANSPORTER 2) (RENAL SODIUM-PHOSPHATE TRANSPORT PROTEIN 2) (RE-
NAL NA+-DEPENDENT PHOSPHATE COTRANSPORTER 2).
1.6e-82:197:72
55 HOMO SAPIENS (HUMAN).
Q06495
- HRIFA010799a

EP 1 130 094 A2

	PROCOLLAGEN ALPHA 1(I) CHAIN PRECURSOR. 1.7e-05:220:30 GALLUS GALLUS (CHICKEN). P02457
5	HRIFA010859a ALPHA-2C-1 ADRENERGIC RECEPTOR (ALPHA-2C-1 ADRENOCEPTOR) (SUBTYPE C4). 0.063:134:33 HOMO SAPIENS (HUMAN). P18825
10	HRIFA010891a HYPOTHETICAL 22.7 KD PROTEIN IN PAS1-MST1 INTERGENIC REGION. 0.044:28:64 SACCHAROMYCES CEREVISIAE (BAKER'S YEAST). P36015
15	HRIFA010975a TYROSINE-PROSINE KINASE SYK (EC 2.7.1.112) (SPLEEN TYROSINE KINASE). 8.5e-113:144:86 HOMO SAPIENS (HUMAN). P43405
20	HRIFA010988a GASTRIN PRECURSOR. 0.084:59:37 HOMO SAPIENS (HUMAN). P01350
25	HRIFA011016a PROTEIN DISULFIDE ISOMERASE-RELATED PROTEIN PRECURSOR (ERP72) (CALCIUM-BINDING PRO- TEIN 2) (CABP2). 3.1e-15:127:37 RATTUS NORVEGICUS (RAT). P38659
30	HRIFA01105a SALIVARY GLUE PROTEIN SGS-7 PRECURSOR. 0.97:41:43 DROSOPHILA MELANOGASTER (FRUIT FLY). P02841
35	HRIFA011105a SALIVARY GLUE PROTEIN SGS-7 PRECURSOR. 0.97:41:43 DROSOPHILA MELANOGASTER (FRUIT FLY). P02841
40	HRIFA011128a GLYCINE-RICH CELL WALL STRUCTURAL PROTEIN (CLONE W10-1) (FRAGMENT). 0.0046:30:63 LYCOPERSICON ESCULENTUM (TOMATO). Q01157
45	HRIFA011179a PROBABLE SERINE/THREONINE-PROTEIN KINASE YKL101W (EC 2.7.1.-). 1.1e-20:127:42 SACCHAROMYCES CEREVISIAE (BAKER'S YEAST). P34244
50	HRIFA011197a DIPEPTIDYL PEPTIDASE IV (EC 3.4.14.5) (DPP IV) (THYMOCYTE-ACTIVATING MOLECULE) (THAM). 5.8e-26:169:40 MUS MUSCULUS (MOUSE).
55	

EP 1 130 094 A2

P28843

HRIFA011449a

GALACTOSYLTRANSFERASE ASSOCIATED PROTEIN KINASE P58/GTA (EC 2.7.1.-).

5 1.9e-26:109:53

MUS MUSCULUS (MOUSE).

P24788

HRIFA011484a

10 D(4) DOPAMINE RECEPTOR (D(2C) DOPAMINE RECEPTOR).

0.00055:115:33

HOMO SAPIENS (HUMAN).

P21917

15 HRIFA011512a

POSSIBLE GLOBAL TRANSCRIPTION ACTIVATOR SNF2L2 (SNF2-ALPHA).

0.00024:139:25

HOMO SAPIENS (HUMAN).

P51531

20

HRIFA011580a

VITELLOGENIN II PRECURSOR (MAJOR VITELLOGENIN) [CONTAINS: LIPOVITELLIN I (LVI) PHOSVITIN (PV) LIPOVITELLIN II (LVII) YGP40].

4.0e-08:182:32

25 GALLUS GALLUS (CHICKEN).

P02845

HRIFA011659a

VON WILLEBRAND FACTOR PRECURSOR.

30 9.8e-17:210:25

HOMO SAPIENS (HUMAN).

P04275

HRIFA011820a

35 ZINC FINGER PROTEIN 136.

1.9e-10:42:73

HOMO SAPIENS (HUMAN).

P52737

40

HRIFA011926a

TRANSCRIPTIONAL REGULATORY PROTEIN ALGP (ALGINATE REGULATORY PROTEIN ALGR3).

1.0:149:22

PSEUDOMONAS AERUGINOSA.

P15276

45

HRIFA011947a

ZINC FINGER PROTEIN 136.

1.3e-80:180:72

HOMO SAPIENS (HUMAN).

50

P52737

HRIFA012069a

A-AGGLUTININ ATTACHMENT SUBUNIT PRECURSOR.

0.0027:205:28

55 SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).

P32323

HRIFA012151a

EP 1 130 094 A2

NEUROGENIC LOCUS NOTCH HOMOLOG PROTEIN 1 PRECURSOR.
0.00028:72:37
RATTUS NORVEGICUS (RAT).
Q07008

5

HRIFA012167a
HYPOTHETICAL SYMPORTER IN GLTS-SELC INTERGENIC REGION.
6.4e-09:145:28
ESCHERICHIA COLI.
10 P31435

HRIFA012278a
ZINC FINGER PROTEIN 140.
3.1e-14:88:52
15 HOMO SAPIENS (HUMAN).
P52738

HRIFA012354a
"SODIUM CHANNEL PROTEIN, BRAIN II ALPHA SUBUNIT."
2.1e-05:120:32
20 RATTUS NORVEGICUS (RAT).
P04775

HRIFA012427a
25 PROCOLLAGEN ALPHA 1(I) CHAIN PRECURSOR.
6.3e-08:250:28
MUS MUSCULUS (MOUSE).
P11087

30 HRIFA012436a
INTESTINAL MEMBRANE A4 PROTEIN (DIFFERENTIATION-DEPENDENT PROTEIN A4).
4.7e-09:95:31
HOMO SAPIENS (HUMAN).
Q04941

35 HRIFA012515a
SODIUM/GLUCOSE COTRANSPORTER 1 (NA+)/GLUCOSE COTRANSPORTER 1) (HIGH AFFINITY SODIUM-GLUCOSE COTRANSPORTER).
3.5e-06:181:27
40 ORYCTOLAGUS CUNICULUS (RABBIT).
P11170

HRIFA012584a
45 MITOCHONDRIAL PRECURSOR PROTEINS IMPORT RECEPTOR (72 KD MITOCHONDRIAL OUTER MEMBRANE PROTEIN) (MITOCHONDRIAL IMPORT RECEPTOR FOR THE ADP/ATP CARRIER) (TRANSLOCASE OF OUTER MEMBRANE TOM70).
4.9e-14:136:29
NEUROSPORA CRASSA.
P23231

50 HRIFA012625a
"HIGH AFFINITY IMMUNOGLOBULIN EPSILON RECEPTOR BETA-SUBUNIT (FCERI) (IGE FC RECEPTOR, BETA-SUBUNIT)."
9.6e-12:103:40
55 RATTUS NORVEGICUS (RAT).
P13386

HRIFA012692a

EP 1 130 094 A2

BLOOM'S SYNDROME PROTEIN.
6.3e-26:203:34
HOMO SAPIENS (HUMAN).
P54132

5

HRIFA012702a
EXTENSIN PRECURSOR (PROLINE-RICH GLYCOPROTEIN).
1.9e-07:153:30
ZEA MAYS (MAIZE).
P14918

10

HRIFA012737a
LEUCOCYTE ANTIGEN CD97 PRECURSOR.
1.6e-09:170:24
HOMO SAPIENS (HUMAN).
P48960

15

HRIFA012795a
VOLTAGE-GATED POTASSIUM CHANNEL PROTEIN KV2.1 (DRK1).
3.0e-34:189:39
RATTUS NORVEGICUS (RAT).
P15387

20

HRIFA012885a
HYPOTHETICAL 30.6 KD PROTEIN IN SCP160-SMC3 INTERGENIC REGION PRECURSOR.
2.9e-21:159:40
SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
P47032

25

HRIFA012914a
ENV POLYPROTEIN PRECURSOR (COAT POLYPROTEIN) [CONTAINS: OUTER MEMBRANE PROTEIN GP70
TRANSMEMBRANE PROTEIN P20E].
3.4e-29:134:47
BABOON ENDOGENOUS VIRUS (STRAIN M7).
P10269

30

HRIFA012969a
ENDOSOMAL P24A PROTEIN PRECURSOR (70 KD ENDOMEMBRANE PROTEIN) (PHEROMONE ALPHA-
FACTOR TRANSPORTER) (ACIDIC 24 KD LATE ENDOCYTIC INTERMEDIATE COMPONENT).
1.2e-30:228:32
SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
P32802

35

HRIFA012990a
PROBABLE CALCIUM-TRANSPORTING ATPASE 6 (EC 3.6.1.38).
7.4e-20:181:33
SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
P39986

40

HRIFA013092a
OUTER MEMBRANE PROTEIN H.8 PRECURSOR.
0.0039:51:39
NEISSERIA GONORRHOEAE.
P11910

45

HRIFA013103a
N-ACETYLLACTOSAMINE SYNTHASE (EC 2.4.1.90) (N-ACETYLGALACTOSAMINE (BETA 1→4)GALACTOSYL-
TRANSFERASE) (EC 2.4.1.38) (LACTOSE SYNTHASE A PROTEIN (EC 2.4.1.22)) (GALACTOSYLTRANS-

50

55

EP 1 130 094 A2

FERASE) (GT).
 0.25:50:34
 MUS MUSCULUS (MOUSE).
 P15535
 5
 HRIFA013135a
 CELL SURFACE GLYCOPROTEIN 1 PRECURSOR (OUTER LAYER PROTEIN B) (SLAYER PROTEIN 1).
 1.6e-05:214:28
 CLOSTRIDIUM THERMOCELLUM.
 10 Q06852
 HRIFA013235a
 PROCOLLAGEN ALPHA 1(III) CHAIN PRECURSOR.
 1.9e-05:113:40
 15 HOMO SAPIENS (HUMAN).
 P02461
 HRIFA013254a
 COMPLEMENT C4 PRECURSOR [CONTAINS: C4A ANAPHYLATOXIN].
 20 3.8e-13:123:41
 MUS MUSCULUS (MOUSE).
 P01029
 HRIFA013265a
 25 CATHEPSIN L PRECURSOR (EC 3.4.22.15) (MAJOR EXCRETED PROTEIN) (MEP).
 7.0e-107:225:86
 HOMO SAPIENS (HUMAN).
 P07711
 30 HRIFA013276a
 5'-NUCLEOTIDASE PRECURSOR (EC 3.1.3.5) (ECTO-NUCLEOTIDASE) (5'-NT) (CD73 ANTIGEN).
 2.2e-117:270:85
 HOMO SAPIENS (HUMAN).
 35 P21589
 HRIFA013279a
 CIRCUMSPOROZOITE PROTEIN PRECURSOR (CS).
 4.9e-05:127:37
 40 PLASMODIUM VIVAX.
 P08677
 HRIFA013376a
 MITOCHONDRIAL PRECURSOR PROTEINS IMPORT RECEPTOR (72 KD MITOCHONDRIAL OUTER MEM-
 45 BRANE PROTEIN) (MITOCHONDRIAL IMPORT RECEPTOR FOR THE ADP/ATP CARRIER) (TRANSLOCASE
 OF OUTER MEMBRANE TOM70).
 8.0e-23:230:31
 NEUROSPORA CRASSA.
 P23231
 50 HRIFA013477a
 OX-2 MEMBRANE GLYCOPROTEIN PRECURSOR (FRAGMENT).
 5.8e-87:197:87
 HOMO SAPIENS (HUMAN).
 55 P41217
 HRIFA013586a
 ENDOZEPINE-RELATED PROTEIN PRECURSOR (MEMBRANE-ASSOCIATED DIAZEPAM BINDING INHIBI-
 TOR) (MA-DBI).

EP 1 130 094 A2

- 3.8e-31:93:64
BOS TAURUS (BOVINE).
P07106
- 5 HRIFA013589a
T-CELL SURFACE PROTEIN TACTILE PRECURSOR (CD96 ANTIGEN).
5.0e-06:95:35
HOMO SAPIENS (HUMAN).
P40200
- 10 HRIFA013620a
"HIGH AFFINITY IMMUNOGLOBULIN EPSILON RECEPTOR BETA-SUBUNIT (FCERI) (IGE FC RECEPTOR, BETA-SUBUNIT)."
7.1e-08:95:37
15 MUS MUSCULUS (MOUSE).
P20490
- HRIFA013726a
SERINE/THREONINE-PROTEIN KINASE STE20 (EC 2.7.1.-).
20 1.5e-33:99:50
SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
Q03497
- HRIFA013744a
25 ENDOZEPINE-RELATED PROTEIN PRECURSOR (MEMBRANE-ASSOCIATED DIAZEPAM BINDING INHIBITOR) (MA-DBI).
7.5e-15:105:38
BOS TAURUS (BOVINE).
P07106
- 30 HRIFA013911a
BIOTINIDASE PRECURSOR (EC 3.5.1.12).
7.8e-37:104:46
HOMO SAPIENS (HUMAN).
35 P43251
- HRIFA013919a
MUCIN 2 PRECURSOR (INTESTINAL MUCIN 2).
1.2e-10:170:32
40 HOMO SAPIENS (HUMAN).
Q02817
- HRIFA013932a
45 "SALIVARY PROLINE-RICH PROTEIN PRECURSOR (CLONES CP3, CP4 AND CP5) [CONTAINS: BASIC PEPTIDE IB-6" PEPTIDE P-H].
2.6e-05:168:34
HOMO SAPIENS (HUMAN).
P04280
- 50 HRIFA013980a
TRANSCRIPTION REGULATORY PROTEIN SNF5 (SWI/SNF COMPLEX COMPONENT SNF5) (TRANSCRIPTION FACTOR TYE4).
0.00036:157:27
SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
55 P18480
- HRIFA014006a
T-CELL SURFACE GLYCOPROTEIN YE1/48 (T LYMPHOCYTE ANTIGEN A1) (LY49-A ANTIGEN).

EP 1 130 094 A2

9.4e-16:185:28
MUS MUSCULUS (MOUSE).
P20937

5 HRIFA014024a
TRANS-ACTING TRANSCRIPTIONAL PROTEIN ICP0 (P135 PROTEIN) (IER 2.9/ER2.6).
0.0013:102:44
BOVINE HERPESVIRUS TYPE 1 (STRAIN JURA).
P29128

10 HRIFA014056a
PROTEIN Q300.
5.1e-05:24:70
MUS MUSCULUS (MOUSE).
15 Q02722

HRIFA014111a
TOLL PROTEIN PRECURSOR.
5.5e-08:203:27
20 DROSOPHILA MELANOGASTER (FRUIT FLY).
P08953

HRIFA014133a
PROLINE-RICH PROTEIN MP-2 PRECURSOR.
25 1.6e-06:143:33
MUS MUSCULUS (MOUSE).
P05142

HRIFA014185a
30 LEUCOCYTE ANTIGEN CD97 PRECURSOR.
6.0e-14:192:30
HOMO SAPIENS (HUMAN).
P48960

35 HRIFA014336a
"GELSOLIN PRECURSOR, PLASMA (ACTIN-DEPOLYMERIZING FACTOR) (ADF) (BREVIN) (FRAGMENT)."
2.8e-70:198:58
SUS SCROFA (PIG).
40 P20305

HRIFA014396a
CREB-BINDING PROTEIN.
2.6e-07:101:34
45 MUS MUSCULUS (MOUSE).
P45481

HRIFA014397a
GENERAL NEGATIVE REGULATOR OF TRANSCRIPTION SUBUNIT 1.
5.2e-05:147:30
50 SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
P25655

HRIFA014465a
55 HYPOTHETICAL 59.1 KD PROTEIN ZK637.1 IN CHROMOSOME III.
2.8e-11:166:30
CAENORHABDITIS ELEGANS.
P30638

EP 1 130 094 A2

5 HRIFA014500a
 HYPOTHETICAL 71.4 KD PROTEIN IN NMD3-ENO2 INTERGENIC REGION.
 1.0e-14:149:35
 SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
 P38862

10 HRIFA014561a
 PROBABLE G PROTEIN-COUPLED RECEPTOR GPR1.
 4.1e-70:156:89
 HOMO SAPIENS (HUMAN).
 P46091

15 HRIFA014568a
 AMINOPEPTIDASE N (EC 3.4.11.2) (MICROSOMAL AMINOPEPTIDASE).
 2.4e-40:196:44
 RATTUS NORVEGICUS (RAT).
 P15684

20 HRIFA014590a
 ATP SYNTHASE PROTEIN 8 (EC 3.6.1.34) (A6L).
 0.18:26:30
 GALLUS GALLUS (CHICKEN).
 P14093

25 HRIFA014598a
 SKIN SECRETORY PROTEIN XP2 PRECURSOR (APEG PROTEIN).
 4.9e-05:124:29
 XENOPUS LAEVIS (AFRICAN CLAWED FROG).
 P17437

30 HRIFA014620a
 ENL PROTEIN.
 0.58:170:30
 HOMO SAPIENS (HUMAN).
 35 Q03111

40 HRIFA014621a
 PREGNANCY-SPECIFIC BETA-1 GLYCOPROTEIN PRECURSOR.
 2.7e-50:150:74
 HOMO SAPIENS (HUMAN).
 P11462

45 HRIFA014688a
 INTEGRIN BETA-6 SUBUNIT PRECURSOR.
 6.9e-31:189:39
 HOMO SAPIENS (HUMAN).
 P18564

50 HRIFA014702a
 PROLINE-RICH PROTEIN MP-2 PRECURSOR.
 6.4e-05:89:40
 MUS MUSCULUS (MOUSE).
 P05142

55 HRIFA014819a
 MICROFIBRIL-ASSOCIATED GLYCOPROTEIN 4.
 7.8e-26:117:46
 HOMO SAPIENS (HUMAN).

EP 1 130 094 A2

	P55083
	HRIFA014868a
5	SALIVARY GLUE PROTEIN SGS-3 PRECURSOR. 8.9e-08:195:29 DROSOPHILA ERECTA (FRUIT FLY). P13730
	HRIFA014951a
10	PLASMINOGEN (EC 3.4.21.7) (FRAGMENT). 4.1e-23:132:39 EQUUS CABALLUS (HORSE). P80010
	HRIFA014967a
15	CHLORINE CHANNEL PROTEIN P64. 2.0e-52:142:76 BOS TAURUS (BOVINE). P35526
20	
	HRIFA015063a
	ZINC FINGER PROTEIN 136. 6.6e-53:229:48 HOMO SAPIENS (HUMAN). P52737
25	
	HRIFA015070a
	SERINE/THREONINE-PROTEIN KINASE NRK1 (EC 2.7.1.-) (N-RICH KINASE 1). 9.3e-24:143:41 SACCHAROMYCES CEREVISIAE (BAKER'S YEAST). P38692
30	
	HRIFA015122a
35	REGULATORY PROTEIN E2. 0.45:129:30 HUMAN PAPILLOMAVIRUS TYPE 5. P06921
	HRIFA015219a
40	FIBRILLIN 1 PRECURSOR (MP340). 9.9e-09:132:32 BOS TAURUS (BOVINE). P98133
	HRIFA015246a
45	PREGNANCY-SPECIFIC BETA-1-GLYCOPROTEIN 4 PRECURSOR (PSBG-4). 2.4e-33:184:46 HOMO SAPIENS (HUMAN). Q00888
50	
	HRIFA015351a
	PROTEOGLYCAN LINK PROTEIN PRECURSOR (CARTILAGE LINK PROTEIN) (LP). 0.0021:122:30 RATTUS NORVEGICUS (RAT). P03994
55	
	HRIFA015423a
	B-CELL LYMPHOMA 3-ENCODED PROTEIN (BCL-3 PROTEIN).

EP 1 130 094 A2

- 1.2e-11:148:35
HOMO SAPIENS (HUMAN).
P20749
- 5 HRIFA015453a
RAC-FAMILY SERINE/THREONINE KINASE HOMOLOG (EC 2.7.1.-).
6.8e-11:91:37
DICTYOSTELIUM DISCOIDEUM (SLIME MOLD).
P54644
- 10 HRIFA015486a
BETA-2-GLYCOPROTEIN I PRECURSOR (APOLIPOPROTEIN H) (APO-H) (ACTIVATED PROTEIN C-BINDING
PROTEIN) (APC INHIBITOR).
2.0e-22:208:27
- 15 MUS MUSCULUS (MOUSE).
Q01339
- HRIFA015506a
COLORECTAL MUTANT CANCER PROTEIN (MCC PROTEIN).
20 1.3e-12:73:50
HOMO SAPIENS (HUMAN).
P23508
- HRIFA015536a
25 CHLORINE CHANNEL PROTEIN P64.
1.2e-49:115:79
BOS TAURUS (BOVINE).
P35526
- HRIFA015547a
30 BEM46 PROTEIN (FRAGMENT).
1.4e-33:137:49
SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).
P54069
- 35 HRIFA015568a
HYPOTHETICAL 35.8 KD PROTEIN C12G12.12 IN CHROMOSOME I.
2.4e-16:152:34
SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).
40 Q09875
- HRIFA015756a
EBNA-2 NUCLEAR PROTEIN.
2.9e-15:28:75
- 45 EPSTEIN-BARR VIRUS (STRAIN B95-8) (HUMAN HERPESVIRUS 4).
P12978
- HRIFA015802a
PROTEOGLYCAN LINK PROTEIN PRECURSOR (CARTILAGE LINK PROTEIN) (LP).
50 0.0035:122:30
RATTUS NORVEGICUS (RAT).
P03994
- HRIFA015811a
55 GLYCOSYLTRANSFERASE ALG2 (EC 2.4.1.-).
6.2e-39:171:43
SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
P43636

EP 1 130 094 A2

5 HRIFA015902a
 A-AGGLUTININ ATTACHMENT SUBUNIT PRECURSOR.
 0.0075:161:29
 SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
 P32323

10 HRIFA015947a
 ZINC FINGER Y-CHROMOSOMAL PROTEIN 1.
 0.035:98:28
 MUS MUSCULUS (MOUSE).
 P10925

15 HRIFA015995a
 PROCOLLAGEN ALPHA 1(III) CHAIN PRECURSOR.
 6.2e-08:221:37
 HOMO SAPIENS (HUMAN).
 P02461

20 HRIFA016070a
 "COMPLEMENT C1Q SUBCOMPONENT, A CHAIN PRECURSOR."
 1.0e-18:179:35
 HOMO SAPIENS (HUMAN).
 P02745

25 HRIFA016214a
 PROLINE-RICH PROTEIN MP-2 PRECURSOR.
 4.0e-05:96:42
 MUS MUSCULUS (MOUSE).
 P05142

30 HRIFA016240a
 HYPOTHETICAL 65.3 KD PROTEIN IN PRE3-SAG1 INTERGENIC REGION.
 8.5e-05:103:33
 SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
 35 P47082

40 HRIFA016255a
 EBNA-1 NUCLEAR PROTEIN.
 4.5e-09:219:33
 EPSTEIN-BARR VIRUS (STRAIN B95-8) (HUMAN HERPESVIRUS 4).
 P03211

45 HRIFA016290a
 COAGULATION FACTOR V PRECURSOR (ACTIVATED PROTEIN C COFACTOR).
 6.7e-21:182:41
 HOMO SAPIENS (HUMAN).
 P12259

50 HRIFA016430a
 ER LUMEN PROTEIN RETAINING RECEPTOR 1 (KDEL RECEPTOR 1).
 7.1e-50:120:86
 HOMO SAPIENS (HUMAN).
 P24390

55 HRIFA016599a
 MEIOTIC RECOMBINATION PROTEIN REC104.
 0.57:73:31
 SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).

EP 1 130 094 A2

P33323

HRIFA016639a

"GLUCOAMYLASE S1/S2 PRECURSOR (EC 3.2.1.3) (GLUCAN 1,4-ALPHA-GLUCOSIDASE) (1,4-ALPHA-D-GLUCAN GLUCOHYDROLASE)."

8.0e-06:206:23

SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).

P08640

HRIFA016654a

HEME-REGULATED EUKARYOTIC INITIATION FACTOR EIF-2-ALPHA KINASE (EC 2.7.1.-) (HRI).

1.1e-78:181:86

ORYCTOLAGUS CUNICULUS (RABBIT).

P33279

HRIFA016669a

ANTER-SPECIFIC PROLINE-RICH PROTEIN APG PRECURSOR.

1.4e-08:87:36

ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).

P40602

HRIFA016758a

GLYCOSYLTRANSFERASE ALG2 (EC 2.4.1.-).

9.5e-17:158:40

SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).

P43636

HRIFA016963a

FMRFAMIDE-RELATED NEUROPEPTIDES PRECURSOR.

6.2e-08:131:32

LYMNAEA STAGNALIS (GREAT POND SNAIL).

P42565

HRIFA017031a

MYOSIN HEAVY CHAIN KINASE A (EC 2.7.1.129) (MHCK A).

2.6e-11:152:34

DICTYOSTELIUM DISCOIDEUM (SLIME MOLD).

P42527

HRIFA017146a

D(4) DOPAMINE RECEPTOR (D(2C) DOPAMINE RECEPTOR).

0.0014:78:37

HOMO SAPIENS (HUMAN).

P21917

HRIFA017190a

FLI-1 ONCOGENE (ERGB TRANSCRIPTION FACTOR).

0.0026:89:30

HOMO SAPIENS (HUMAN).

Q01543

HRIFA017257a

"GELSOLIN PRECURSOR, PLASMA (ACTIN-DEPOLYMERIZING FACTOR) (ADF) (BREVIN) (AGEL)."

2.5e-79:261:57

HOMO SAPIENS (HUMAN).

P06396

HRIFA017295a

EP 1 130 094 A2

*ALPHA-1,6-MANNOSYL-GLYCOPROTEIN BETA-1,2-N-ACETYLGLUCOSAMINYLTRANSFERASE (EC 2.4.1.143) (N-GLYCOSYL-OLIGOSACCHARIDE-GLYCOPROTEIN N-ACETYLGLUCOSAMINYLTRANSFERASE II) (BETA-1,2-N-ACETYLGLUCOSAMINYLTRANSFERASE II) (GNT-II) (GLCNAC-T II)."
3.4e-20:66:78
5 HOMO SAPIENS (HUMAN).
Q10469

HRIFA017312a
10 C4B-BINDING PROTEIN ALPHA CHAIN PRECURSOR (PROLINE-RICH PROTEIN) (PRP).
2.7e-19:221:33
HOMO SAPIENS (HUMAN).
P04003

HRIFA017456a
15 LAMININ ALPHA-1 CHAIN PRECURSOR (LAMININ A CHAIN).
0.11:94:35
MUS MUSCULUS (MOUSE).
P19137

HRIFA017457a
20 SYNAPTOTAGMIN II.
7.2e-07:98:35
MUS MUSCULUS (MOUSE).
P46097

HRIFA017643a
25 NOV PROTEIN HOMOLOG PRECURSOR (NOVH).
2.2e-07:81:41
HOMO SAPIENS (HUMAN).
30 P48745

HRIFA017670a
TRANS-GOLGI NETWORK INTEGRAL MEMBRANE PROTEIN TGN38 PRECURSOR.
4.9e-06:172:27
35 RATTUS NORVEGICUS (RAT).
P19814

HRIFA017703a
40 PUTATIVE SERINE/THREONINE-PROTEIN KINASE PKWA (EC 2.7.1.-).
1.9e-16:129:34
THERMOMONOSPORA CURVATA.
P49695

HRIFA017791a
45 MUCIN 2 PRECURSOR (INTESTINAL MUCIN 2).
0.012:71:38
HOMO SAPIENS (HUMAN).
Q02817

HRIFA017801a
50 PROLINE-RICH PROTEIN MP-2 PRECURSOR.
4.5e-07:86:39
MUS MUSCULUS (MOUSE).
P05142

HRIFA017818a
55 ATP SYNTHASE C CHAIN (EC 3.6.1.34) (LIPID-BINDING PROTEIN).
1.0:32:40

EP 1 130 094 A2

STREPTOMYCES LIVIDANS.

P50014

HRIFA017836a

5 "TETRACYCLINE RESISTANCE PROTEIN, CLASS H (TETA(H))."

1.3e-08:113:31

PASTEURELLA MULTOCIDA.

P51564

10

HRIFA017855a

ORM1 PROTEIN.

1.7e-18:137:35

SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).

P53224

15

HRIFA017921a

TRANS-ACTING TRANSCRIPTIONAL PROTEIN ICP0 (IMMEDIATE-EARLY PROTEIN IE110) (VMW110) (ALPHA-0 PROTEIN).

2.0e-09:182:35

20

HERPES SIMPLEX VIRUS (TYPE 1 / STRAIN 17).

P08393

HRIFA018092a

25 "DIACYLGLYCEROL CHOLINEPHOSPHOTRANSFERASE (EC 2.7.8.2) (SN-1,2-DIACYLGLYCEROL CHOLINEPHOSPHOTRANSFERASE) (CHOPT)."

2.1e-20:119:42

SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).

P17898

30

HRIFA018131a

ORM1 PROTEIN.

2.6e-20:137:37

SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).

P53224

35

HRIFA018134a

SERINE/THREONINE-PROTEIN KINASE CTR1 (EC 2.7.1.37).

1.1e-11:147:32

ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).

40

Q05609

HRIFA018238a

NEUROGENIC LOCUS NOTCH HOMOLOG PROTEIN 1 PRECURSOR (MOTCH PROTEIN).

8.6e-06:74:44

45

MUS MUSCULUS (MOUSE).

Q01705

HRIFA018262a

PUTATIVE SERINE/THREONINE-PROTEIN KINASE PKWA (EC 2.7.1.-).

50

6.4e-10:71:38

THERMOMONOSPORA CURVATA.

P49695

HRIFA018287a

55 HYPOTHETICAL 57.5 KD PROTEIN IN VMA7-RPS25A INTERGENIC REGION.

1.5e-06:214:32

SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).

P53214

EP 1 130 094 A2

HRIFA018447a
SKIN SECRETORY PROTEIN XP2 PRECURSOR (APEG PROTEIN).
0.00065:133:33
XENOPUS LAEVIS (AFRICAN CLAWED FROG).
5 P17437

HRIFA018580a
COP-COATED VESICLE MEMBRANE PROTEIN P24 PRECURSOR (FRAGMENT).
2.1e-18:109:41
10 CRICETULUS GRISEUS (CHINESE HAMSTER).
P49020

HRIFA018666a
PROTEIN-TYROSINE PHOSPHATASE DLAR PRECURSOR (EC 3.1.3.48) (PROTEIN-TYROSINE-PHOSPHATE
15 PHOSPHOHYDROLASE).
1.7e-06:191:28
DROSOPHILA MELANOGASTER (FRUIT FLY).
P16621

HRIFA018688a
20 PHLB PROTEIN PRECURSOR.
1.9e-06:110:35
SERRATIA LIQUEFACIENS.
P18954

HRIFA018754a
25 "GLUCOAMYLASE S1/S2 PRECURSOR (EC 3.2.1.3) (GLUCAN 1,4-ALPHA-GLUCOSIDASE) (1,4-ALPHA-D-
GLUCAN GLUCOHYDROLASE)."
1.8e-06:195:27
30 SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
P08640

HRIFA018794a
MSP1 PROTEIN HOMOLOG.
35 3.2e-06:93:25
CAENORHABDITIS ELEGANS.
P54815

HRIFA018827a
40 HYPOTHETICAL 59.1 KD PROTEIN ZK637.1 IN CHROMOSOME III.
3.1e-17:180:28
CAENORHABDITIS ELEGANS.
P30638

HRIFA018870a
45 HYPOTHETICAL 35.6 KD PROTEIN IN SPC1-ILV3 INTERGENIC REGION.
4.7e-09:70:37
SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
P47088

HRIFA018904a
50 MYOTONIN-PROTEIN KINASE (EC 2.7.1.-) (MYOTONIC DYSTROPHY PROTEIN KINASE) (MDPK) (DM-KI-
NASE) (DMK) (DMPK) (MT-PK).
5.5e-12:142:32
55 HOMO SAPIENS (HUMAN).
Q09013

HRIFA018931a

EP 1 130 094 A2

- ZINC FINGER PROTEIN 140.
2.9e-10:47:74
HOMO SAPIENS (HUMAN).
P52738
- 5
HRIFA018993a
HYPOTHETICAL 21.5 KD PROTEIN IN SEC15-SAP4 INTERGENIC REGION.
1.2e-13:117:34
SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
10 P53073
- HRIFA019105a
DORSAL-VENTRAL PATTERNING TOLLOID PROTEIN PRECURSOR (EC 3.4.24.-).
7.5e-22:203:32
15 DROSOPHILA MELANOGASTER (FRUIT FLY).
P25723
- HRIFA019136a
"MYRISTOYLATED ALANINE-RICH C-KINASE SUBSTRATE (MARCKS) (PROTEIN KINASE C SUBSTRATE,
20 80 KD PROTEIN, LIGHT CHAIN) (PKCSL) (80K-L PROTEIN)."
1.0e-25:74:81
HOMO SAPIENS (HUMAN).
P29966
- 25 HRIFA019175a
PROTEIN KINASE WIS1 (EC 2.7.1.-).
1.3e-14:84:39
SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).
30 P33886
- HRIFA019262a
ZINC FINGER PROTEIN 85 (ZINC FINGER PROTEIN HPF4) (HTF1).
2.5e-55:188:50
HOMO SAPIENS (HUMAN).
35 Q03923
- HRIFA019412a
CATHEPSIN E PRECURSOR (EC 3.4.23.34).
1.4e-09:121:33
40 CAVIA PORCELLUS (GUINEA PIG).
P25796
- HRIFA019437a
REGULATORY PROTEIN E2.
45 0.26:77:37
HUMAN PAPILLOMAVIRUS TYPE 14.
P36783
- HRIFA019466a
50 EBNA-1 NUCLEAR PROTEIN.
2.7e-19:130:43
EPSTEIN-BARR VIRUS (STRAIN B95-8) (HUMAN HERPESVIRUS 4).
P03211
- 55 HRIFA019490a
TRANSCRIPTION REGULATORY PROTEIN SNF5 (SWI/SNF COMPLEX COMPONENT SNF5) (TRANSCRIPTION FACTOR TYE4).
1.1e-09:132:34

EP 1 130 094 A2

SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
P18480

5 HRIFA019498a
VOLTAGE-GATED POTASSIUM CHANNEL PROTEIN SHAL (SHAL2).
5.6e-05:87:36
DROSOPHILA MELANOGASTER (FRUIT FLY).
P17971

10 HRIFA019532a
EBNA-1 NUCLEAR PROTEIN.
1.8e-05:67:49
EPSTEIN-BARR VIRUS (STRAIN B95-8) (HUMAN HERPESVIRUS 4).
P03211

15 HRIFA019651a
ACIDIC PHOSPHOPROTEIN PRECURSOR (50 KD ANTIGEN).
6.1e-05:31:64
PLASMODIUM CHABAUDI.
20 Q02752

HRIFA019867a
RENAL SODIUM-DEPENDENT PHOSPHATE TRANSPORT PROTEIN 2 (SODIUM/PHOSPHATE COTRANS-
PORTER 2) (NA(+)/PI COTRANSPORTER 2) (RENAL SODIUM-PHOSPHATE TRANSPORT PROTEIN 2) (RE-
25 NAL NA+-DEPENDENT PHOSPHATE COTRANSPORTER 2).
8.2e-34:103:71
RATTUS NORVEGICUS (RAT).
Q06496

30 HRIFA019869a
SERINE/THREONINE-PROTEIN KINASE FUSED (EC 2.7.1.-).
7.2e-29:83:49
DROSOPHILA MELANOGASTER (FRUIT FLY).
P23647

35 HRIFA019958a
REPRESSOR PROTEIN CI (FRAGMENT).
0.99:45:37
BACTERIOPHAGE 434.
40 P16117

HRIFA020144a
SKIN SECRETORY PROTEIN XP2 PRECURSOR (APEG PROTEIN).
2.8e-06:176:30
45 XENOPUS LAEVIS (AFRICAN CLAWED FROG).
P17437

HRIFA020184a
NON-RECEPTOR TYROSINE KINASE SPORE LYSIS A (EC 2.7.1.112) (TYROSINE-PROTEIN KINASE 1).
50 1.9e-10:102:37
DICTYOSTELIUM DISCOIDEUM (SLIME MOLD).
P18160

HRIFA020272a
55 MUSCARINIC ACETYLCHOLINE RECEPTOR M3.
5.5e-91:211:85
HOMO SAPIENS (HUMAN).
P20309

EP 1 130 094 A2

HRIFA020335a
 PUTATIVE SERINE/THREONINE-PROTEIN KINASE P78 (EC 2.7.1.-).
 5.0e-104:275:72
 HOMO SAPIENS (HUMAN).
 5 P27448

HRIFA020349a
 BRITTLE-1 PROTEIN PRECURSOR.
 6.0e-30:214:35
 10 ZEA MAYS (MAIZE).
 P29518

HRIFA020453a
 PROTEIN TRANSPORT PROTEIN SEC22 (PROTEIN SLY2).
 15 2.5e-08:132:28
 SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
 P22214

HRIFA020693a
 20 TUMOR SUPPRESSOR PROTEIN DCC PRECURSOR (COLORECTAL CANCER SUPPRESSOR).
 3.9e-09:96:35
 HOMO SAPIENS (HUMAN).
 P43146

HRIFA020707a
 25 PROCYCLIC FORM SPECIFIC POLYPEPTIDE B-ALPHA PRECURSOR (PROCYCLIN) (PARP B-ALPHA) (PS-
 SA-1).
 3.4e-09:95:33
 TRYPANOSOMA BRUCEI BRUCEI.
 30 Q06084

HRIFA020748a
 EXTENSIN PRECURSOR (CELL WALL HYDROXYPROLINE-RICH GLYCOPROTEIN).
 3.2e-09:210:28
 35 NICOTIANA TABACUM (COMMON TOBACCO).
 P13983

HRIFA020862a
 MODIFIER 3 PROTEIN (M33).
 40 5.6e-26:76:61
 MUS MUSCULUS (MOUSE).
 P30658

HRIFA020883a
 45 PROTEIN Q300.
 0.00054:21:66
 MUS MUSCULUS (MOUSE).
 Q02722

HRIFA021007a
 50 TRANS-ACTING TRANSCRIPTIONAL PROTEIN ICP0 (IMMEDIATE-EARLY PROTEIN IE110) (VMW110) (AL-
 PHA-0 PROTEIN).
 0.092:73:36
 HERPES SIMPLEX VIRUS (TYPE 1 / STRAIN 17).
 55 P08393

HRIFA021040a
 TRANSCRIPTION FACTOR GATA-4 (GATA BINDING FACTOR-4).

EP 1 130 094 A2

- 0.98:63:39
HOMO SAPIENS (HUMAN).
P43694
- 5 HRIFA021061a
SKIN SECRETORY PROTEIN XP2 PRECURSOR (APEG PROTEIN).
2.8e-09:162:31
XENOPUS LAEVIS (AFRICAN CLAWED FROG).
P17437
- 10 HRIFA021213a
OLIGOSACCHARYL TRANSFERASE STT3 SUBUNIT HOMOLOG.
2.0e-38:96:72
CAENORHABDITIS ELEGANS.
15 P46975
- HRIFA021224a
RENAL TRANSCRIPTION FACTOR KID-1 (TRANSCRIPTION FACTOR 17).
2.8e-06:55:52
20 RATTUS NORVEGICUS (RAT).
Q02975
- HRIFA021398a
COAGULATION FACTOR VII PRECURSOR (EC 3.4.21.21).
2.5e-17:78:51
25 ORYCTOLAGUS CUNICULUS (RABBIT).
P98139
- HRIFA021445a
30 PRE-B-CELL LEUKEMIA TRANSCRIPTION FACTOR-1 (HOMEBOX PROTEIN PBX1) (HOMEBOX PROTEIN
PRL).
0.38:146:31
HOMO SAPIENS (HUMAN).
P40424
- 35 HRIFA021494a
EBNA-1 NUCLEAR PROTEIN.
6.8e-07:116:41
EPSTEIN-BARR VIRUS (STRAIN B95-8) (HUMAN HERPESVIRUS 4).
40 P03211
- HRIFA021499a
CARTILAGE MATRIX PROTEIN PRECURSOR (MATRILIN-1).
7.1e-34:159:50
45 GALLUS GALLUS (CHICKEN).
P05099
- HRIFA021543a
ANTITHROMBIN-III PRECURSOR (ATIII) (FRAGMENT).
50 0.0087:50:40
GALLUS GALLUS (CHICKEN).
Q03352
- HRIFA021620a
55 PLATELET FACTOR 4 (PF-4).
0.019:65:27
SUS SCROFA (PIG).
P30034

EP 1 130 094 A2

5 HRIFA021637a
 CARTILAGE MATRIX PROTEIN PRECURSOR (MATRILIN-1).
 6.0e-37:147:53
 GALLUS GALLUS (CHICKEN).
 P05099

10 HRIFA021651a
 CARG-BINDING FACTOR-A (CBF-A).
 2.6e-11:170:30
 MUS MUSCULUS (MOUSE).
 Q99020

15 HRIFA021754a
 CARTILAGE MATRIX PROTEIN PRECURSOR (MATRILIN-1).
 1.2e-37:137:51
 GALLUS GALLUS (CHICKEN).
 P05099

20 HRIFA021781a
 DNA-REPAIR PROTEIN COMPLEMENTING XP-D CELLS (XERODERMA PIGMENTOSUM GROUP D COM-
 PLEMENTING PROTEIN) (DNA EXCISION REPAIR PROTEIN ERCC-2).
 7.1e-19:199:31
 HOMO SAPIENS (HUMAN).
 P18074

25 HRIFA021787a
 PROTEIN Q300.
 0.051:13:84
 MUS MUSCULUS (MOUSE).
 30 Q02722

35 HRIFA021794a
 RAC-FAMILY SERINE/THREONINE KINASE HOMOLOG (EC 2.7.1.-).
 1.6e-07:90:32
 DICTYOSTELIUM DISCOIDEUM (SLIME MOLD).
 P54644

40 HRIFA021855a
 SKIN SECRETORY PROTEIN XP2 PRECURSOR (APEG PROTEIN).
 8.6e-06:163:30
 XENOPUS LAEVIS (AFRICAN CLAWED FROG).
 P17437

45 HRIFA021906a
 S-ANTIGEN PROTEIN PRECURSOR.
 2.1e-09:226:28
 PLASMODIUM FALCIPARUM (ISOLATE V1).
 P09593

50 HRIFA022055a
 BETA-LYTIC METALLOENDOPEPTIDASE PRECURSOR (EC 3.4.24.32) (BETA-LYTIC PROTEASE).
 0.63:118:31
 ACHROMOBACTER LYTICUS.
 P27458

55 HRIFA022065a
 BETA GALACTOSIDASE-RELATED PROTEIN PRECURSOR.
 9.7e-24:235:34

EP 1 130 094 A2

- HOMO SAPIENS (HUMAN).
P16279
- 5 HRIFA022139a
HYPOTHETICAL 85.7 KD PROTEIN C13G6.03 IN CHROMOSOME I.
2.1e-57:232:52
SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).
Q09782
- 10 HRIFA022156a
GLUTENIN, HIGH MOLECULAR WEIGHT SUBUNIT PW212 PRECURSOR.
1.4e-07:133:35
TRITICUM AESTIVUM (WHEAT).
P08489
- 15 HRIFA022166a
EXCISION REPAIR PROTEIN ERCC-6 (COCKAYNE SYNDROME PROTEIN CSB).
3.5e-28:194:35
HOMO SAPIENS (HUMAN).
20 Q03468
- HRIFA022177a
PUTATIVE SERINE/THREONINE-PROTEIN KINASE PKWA (EC 2.7.1.-).
2.2e-12:137:32
25 THERMOMONOSPORA CURVATA.
P49695
- HRIFA022182a
SERINE/THREONINE-PROTEIN KINASE MAK (EC 2.7.1.-) (MALE GERM CELL-ASSOCIATED KINASE).
30 1.2e-47:121:79
RATTUS NORVEGICUS (RAT).
P20793
- HRIFA022203a
35 COLLAGEN ALPHA 1 (III) CHAIN.
1.1e-05:211:33
BOS TAURUS (BOVINE).
P04258
- 40 HRIFA022227a
POTENTIAL CAAX PRENYL PROTEASE 1 (EC 3.4.24.-) (PRENYL PROTEIN- SPECIFIC ENDOPROTEASE 1)
(PPSEP 1).
3.2e-31:229:36
SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).
45 Q10071
- HRIFA022234a
CARBOXYPEPTIDASE KEX1 PRECURSOR (EC 3.4.16.6) (CARBOXYPEPTIDASE D).
1.8e-08:110:30
50 SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
P09620
- HRIFA022249a
55 ZINC FINGER PROTEIN 133.
1.1e-34:84:48
HOMO SAPIENS (HUMAN).
P52736

EP 1 130 094 A2

- HRIFA022265a
CALCIUM/CALMODULIN-DEPENDENT PROTEIN KINASE TYPE IV CATALYTIC CHAIN (EC 2.7.1.123) (CAM
KINASE-GR) (CAMK IV) [CONTAINS: CALSPERMIN].
5.1e-26:188:40
5 RATTUS NORVEGICUS (RAT).
P13234
- HRIFA022328a
SCO1 PROTEIN PRECURSOR.
10 5.4e-25:84:45
SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
P23833
- HRIFA022335a
15 TRANS-ACTING TRANSCRIPTIONAL PROTEIN ICP0 (IMMEDIATE-EARLY PROTEIN IE110) (VMW110) (AL-
PHA-0 PROTEIN).
0.21:121:29
HERPES SIMPLEX VIRUS (TYPE 1 / STRAIN 17).
P08393
20
- HRIFA022348a
AGAMOUS PROTEIN.
1.0:40:42
ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).
25 P17839
- HRIFA022411a
TRANS-ACTING TRANSCRIPTIONAL PROTEIN ICP0 (IMMEDIATE-EARLY PROTEIN IE110) (VMW110) (AL-
PHA-0 PROTEIN).
30 0.00059:111:35
HERPES SIMPLEX VIRUS (TYPE 1 / STRAIN 17).
P08393
- HRIFA022423a
35 HYPOTHETICAL 24.5 KD PROTEIN IN SAP185-BCK1 INTERGENIC REGION.
2.5e-15:106:42
SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
P40857
- HRIFA022462a
40 RETINOIC ACID RECEPTOR RXR-BETA.
0.0010:124:33
HOMO SAPIENS (HUMAN).
P28702
45
- HRIFA022493a
ALPHA-2A ADRENERGIC RECEPTOR (ALPHA-2A ADRENOCEPTOR) (ALPHA-2AAR).
0.0018:130:34
MUS MUSCULUS (MOUSE).
50 Q01338
- HRIFA022528a
EXTENSIN PRECURSOR (PROLINE-RICH GLYCOPROTEIN).
3.2e-23:230:28
55 ZEA MAYS (MAIZE).
P14918
- HRIFA022546a

EP 1 130 094 A2

	NINAC SHORT PROTEIN (EC 2.7.1.-).
	8.5e-42:209:43
	DROSOPHILA MELANOGASTER (FRUIT FLY).
	P10677
5	HRIFA022564a
	ZINC FINGER PROTEIN 140.
	7.9e-23:116:51
	HOMO SAPIENS (HUMAN).
10	P52738
	HRIFA022616a
	LOW-DENSITY LIPOPROTEIN RECEPTOR-RELATED PROTEIN 1 PRECURSOR (LRP) (ALPHA-2-MAC-
	ROGLOBULIN RECEPTOR) (A2MR) (APOLIPOPROTEIN E RECEPTOR) (APOER) (CD91).
15	7.4e-36:172:43
	HOMO SAPIENS (HUMAN).
	Q07954
	HRIFA022671a
20	PAIRED AMPHIPATHIC HELIX PROTEIN.
	2.0e-26:186:36
	SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
	P22579
	HRIFA022691a
25	FIBRINOGEN-LIKE PROTEIN A PRECURSOR (FREPE-A).
	1.4e-44:229:41
	PARASTICHOPUS PARVIMENSIS (SEA CUCUMBER).
	P19477
30	HRIFA022702a
	PROCOLLAGEN ALPHA 1(I) CHAIN PRECURSOR.
	1.1e-08:146:38
	GALLUS GALLUS (CHICKEN).
35	P02457
	HRIFA022707a
	GC-RICH SEQUENCE DNA-BINDING FACTOR (GCF) (TRANSCRIPTION FACTOR 9) (TCF-9).
	7.0e-40:229:37
40	HOMO SAPIENS (HUMAN).
	P16383
	HRIFA022714a
	"AMELOGENIN, CLASS I PRECURSOR."
45	0.62:96:31
	BOS TAURUS (BOVINE).
	P02817
	HRIFA022728a
50	ACROSIN PRECURSOR (EC 3.4.21.10) (53 KD FUCOSE-BINDING PROTEIN).
	1.7e-06:28:64
	SUS SCROFA (PIG).
	P08001
55	HRIFA022729a
	"ALPHA-1,6-MANNOSYL-GLYCOPROTEIN BETA-1,2-N-ACETYLGLUCOSAMINYLTRANSFERASE (EC
	2.4.1.143) (N-GLYCOSYL-OLIGOSACCHARIDE-GLYCOPROTEIN N-ACETYLGLUCOSAMINYLTRANS-
	FERASE II) (BETA-1,2-N-ACETYLGLUCOSAMINYLTRANSFERASE II) (GNT-II) (GLCNAC-T II)."

EP 1 130 094 A2

- 7.7e-29:69:84
HOMO SAPIENS (HUMAN).
Q10469
- 5 HRIFA022737a
TENASCIN PRECURSOR (TN) (HEXABRACHION) (CYTOTACTIN) (NEURONECTIN) (GMEM) (JI) (MIOTEND-
INOUS ANTIGEN) (GLIOMA-ASSOCIATED-EXTRACELLULAR MATRIX ANTIGEN) (GP 150-225).
6.7e-19:170:37
GALLUS GALLUS (CHICKEN).
10 P10039
- HRIFA022776a
PROBABLE PROTEIN DISULFIDE ISOMERASE P5 PRECURSOR (EC 5.3.4.1).
4.0e-20:199:31
15 MEDICAGO SATIVA (ALFALFA).
P38661
- HRIFA022782a
CIRCUMSPOROZOITE PROTEIN PRECURSOR (CS).
20 3.7e-09:184:36
PLASMODIUM CYNOMOLGI (STRAIN BEROK).
P08672
- HRIFA022865a
25 COLLAGEN ALPHA 1(III) CHAIN.
2.5e-09:169:33
BOS TAURUS (BOVINE).
P04258
- HRIFA022875a
30 BONE PROTEOGLYCAN II PRECURSOR (PG-S2) (DECORIN).
9.1e-14:115:33
BOS TAURUS (BOVINE).
P21793
- 35 HRIFA022890a
SKIN SECRETORY PROTEIN XP2 PRECURSOR (APEG PROTEIN).
1.8e-10:237:30
XENOPUS LAEVIS (AFRICAN CLAWED FROG).
40 P17437
- HRIFA022895a
ZINC FINGER PROTEIN 85 (ZINC FINGER PROTEIN HPF4) (HTF1).
2.4e-106:283:67
45 HOMO SAPIENS (HUMAN).
Q03923
- HRIFA022985a
50 PROCYCLIC FORM SPECIFIC POLYPEPTIDE B-ALPHA PRECURSOR (PROCYCLIN) (PARP B-ALPHA) (PS-
SA-1).
3.0e-10:33:72
TRYPANOSOMA BRUCEI BRUCEI.
Q06084
- 55 HRIFA023007a
MHC CLASS II REGULATORY FACTOR RFX1 (RFX) (ENHANCER FACTOR C) (EF-C).
1.1e-27:66:54
HOMO SAPIENS (HUMAN).

EP 1 130 094 A2

P22670

- HRIFA023048a
COLLAGEN ALPHA 1(I) CHAIN (FRAGMENTS).
5 2.2e-07:221:33
RATTUS NORVEGICUS (RAT).
P02454
- HRIFA023069a
10 BASEMENT MEMBRANE-SPECIFIC HEPARAN SULFATE PROTEOGLYCAN CORE PROTEIN PRECURSOR
(HSPG) (PERLECAN) (PLC).
3.4e-08:149:31
HOMO SAPIENS (HUMAN).
P98160
- HRIFA023129a
15 HISTIDINE-RICH GLYCOPROTEIN PRECURSOR.
4.2e-06:37:51
PLASMODIUM LOPHURAE.
20 P04929
- HRIFA023154a
GLYCOPROTEIN X PRECURSOR.
8.2e-05:140:27
25 EQUINE HERPESVIRUS TYPE 1 (STRAIN AB4P) (EHV-1).
P28968
- HRIFA023212a
30 A-AGGLUTININ ATTACHMENT SUBUNIT PRECURSOR.
8.3e-10:249:32
SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
P32323
- HRIFA023227a
35 GALACTOSE-PROTON SYMPORT (GALACTOSE TRANSPORTER).
9.2e-15:180:30
ESCHERICHIA COLI.
P37021
- HRIFA023257a
40 PROTEIN TRANSPORT PROTEIN SEC61 ALPHA SUBUNIT.
2.4e-118:229:88
RATTUS NORVEGICUS (RAT).
P38378
- HRIFA023304a
45 PROBABLE CALCIUM-TRANSPORTING ATPASE 3 (EC 3.6.1.38) (ENDOPLASMIC RETICULUM CA2+-AT-
PASE).
1.3e-23:222:29
50 SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
P39524
- HRIFA023434a
55 VOLTAGE-GATED POTASSIUM CHANNEL PROTEIN KV1.6 (RCK2) (KV2).
0.00018:157:30
RATTUS NORVEGICUS (RAT).
P17659

EP 1 130 094 A2

- HRIFA023464a
GLYCINE-RICH CELL WALL STRUCTURAL PROTEIN 2 PRECURSOR.
1.0e-11:75:46
ORYZA SATIVA (RICE).
5 P29834
- HRIFA023489a
HYPOTHETICAL 111.9 KD PROTEIN C22H10.03C IN CHROMOSOME I.
4.4e-09:230:23
10 SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).
Q10297
- HRIFA023634a
EBNA-1 NUCLEAR PROTEIN.
15 1.8e-08:113:45
EPSTEIN-BARR VIRUS (STRAIN B95-8) (HUMAN HERPESVIRUS 4).
P03211
- HRIFA023767a
20 CYTOCHROME B5.
1.1e-12:92:38
SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
P40312
- HRIFA023894a
25 PROLINE-RICH PROTEIN MP-2 PRECURSOR.
3.6e-05:80:40
MUS MUSCULUS (MOUSE).
P05142
30
- HRIFA023923a
CYTOCHROME C OXIDASE POLYPEPTIDE I (EC 1.9.3.1).
4.2e-76:128:85
HOMO SAPIENS (HUMAN).
35 P00395
- HRIFA024088a
NEURON-SPECIFIC X11 PROTEIN (FRAGMENT).
1.1e-05:118:32
40 MUS MUSCULUS (MOUSE).
P98084
- HRIFA024132a
45 VOLTAGE-GATED POTASSIUM CHANNEL PROTEIN KV1.9.
6.5e-40:136:61
HOMO SAPIENS (HUMAN).
P51787
- HRIFA024185a
50 ETS-DOMAIN TRANSCRIPTION FACTOR ERF.
0.55:128:29
HOMO SAPIENS (HUMAN).
P50548
- HRIFA024197a
55 MITOCHONDRIAL PRECURSOR PROTEINS IMPORT RECEPTOR (72 KD MITOCHONDRIAL OUTER MEM-
BRANE PROTEIN) (MITOCHONDRIAL IMPORT RECEPTOR FOR THE ADP/ATP (CARRIER) (TRANSLOCASE
OF OUTER MEMBRANE TOM70).

EP 1 130 094 A2

- 7.5e-09:93:34
NEUROSPORA CRASSA.
P23231
- 5 HRIFA024218a
PROCOLLAGEN ALPHA 1(I) CHAIN PRECURSOR.
6.7e-06:180:36
HOMO SAPIENS (HUMAN).
P02452
- 10 HRIFA024255a
HYPOTHETICAL 116.3 KD PROTEIN C26F1.09 IN CHROMOSOME I.
4.8e-23:172:33
SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).
15 Q10496
- HRIFA024305a
CYCLIC-AMP-DEPENDENT TRANSCRIPTION FACTOR ATF-6 (FRAGMENT).
0.047:47:29
20 HOMO SAPIENS (HUMAN).
P18850
- HRIFA024392a
TRANSMEMBRANE PROTEIN SEX PRECURSOR.
25 6.7e-24:119:43
HOMO SAPIENS (HUMAN).
P51805
- HRIFA024423a
30 COP-COATED VESICLE MEMBRANE PROTEIN P24 PRECURSOR (FRAGMENT).
2.1e-18:109:41
CRICETULUS GRISEUS (CHINESE HAMSTER).
P49020
- 35 HRIFA024473a
COLLAGEN ALPHA 1(I) CHAIN (FRAGMENTS).
3.3e-05:106:41
BOS TAURUS (BOVINE).
P02453
- 40 HRIFA024482a
PSTIL-SPECIFIC EXTENSIN-LIKE PROTEIN PRECURSOR (PELP).
1.2e-07:99:31
NICOTIANA TABACUM (COMMON TOBACCO).
45 Q03211
- HRIFA024504a
ESTRADIOL 17 BETA-DEHYDROGENASE 3 (EC 1.1.1.62) (17-BETA-HSD 3) (TESTICULAR 17-BETA-HY-
50 DROXYSTEROID DEHYDROGENASE).
2.6e-43:205:49
HOMO SAPIENS (HUMAN).
P37058
- 55 HRIFA024543a
GLYCOPROTEIN X PRECURSOR.
1.5e-06:257:28
EQUINE HERPESVIRUS TYPE 1 (STRAIN AB4P) (EHV-1).
P28968

EP 1 130 094 A2

- HRIFA024718a
BETA-GALACTOSIDASE PRECURSOR (EC 3.2.1.23) (LACTASE) (ACID BETA-GALACTOSIDASE).
5.3e-45:168:52
MUS MUSCULUS (MOUSE).
5 P23780
- HRIFA024767a
SODIUM CHANNEL PROTEIN (NA⁺ CHANNEL).
7.4e-30:221:31
10 ELECTROPHORUS ELECTRICUS (ELECTRIC EEL).
P02719
- HRIFA024884a
HYPOTHETICAL 13.6 KD PROTEIN IN SPT4-ROM1 INTERGENIC REGION.
15 0.0089:23:65
SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
P53245
- HRIFA024893a
20 REGULATORY PROTEIN E2.
0.0021:167:31
HUMAN PAPILLOMAVIRUS TYPE 8.
P06422
- HRIFA024937a
25 GNS1 PROTEIN.
1.0e-15:173:33
SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
P25358
30
- HRIFA024978a
MUCIN 2 PRECURSOR (INTESTINAL MUCIN 2).
0.00019:150:32
HOMO SAPIENS (HUMAN).
35 Q02817
- HRIFA024994a
EXTENSIN PRECURSOR (CELL WALL HYDROXYPROLINE-RICH GLYCOPROTEIN).
5.3e-22:145:46
40 NICOTIANA TABACUM (COMMON TOBACCO).
P13983
- HRIFA025033a
TRANS-ACTING TRANSCRIPTIONAL PROTEIN ICP0 (VMW118 PROTEIN).
45 0.50:215:29
HERPES SIMPLEX VIRUS (TYPE 2 / STRAIN HG52).
P28284
- HRIFA025046a
50 PROBABLE CALCIUM-TRANSPORTING ATPASE 6 (EC 3.6.1.38).
1.7e-41:104:48
SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
P39986
- HRIFA025250a
55 *PROTEIN KINASE C, BRAIN ISOZYME (EC 2.7.1.-) (PKC) (DPKC53E(BR)).*
7.4e-17:126:34
DROSOPHILA MELANOGASTER (FRUIT FLY).

EP 1 130 094 A2

P05130

HRIFA025261a
MYOSIN I ALPHA (MMI-ALPHA).
2.3e-64:141:84
MUS MUSCULUS (MOUSE).
P46735

5

HRIFA025290a
EBNA-1 NUCLEAR PROTEIN.
0.016:79:40
EPSTEIN-BARR VIRUS (STRAIN B95-8) (HUMAN HERPESVIRUS 4).
P03211

10

HRIFA025327a
PROLINE-RICH PROTEIN MP-2 PRECURSOR.
2.3e-06:104:37
MUS MUSCULUS (MOUSE).
P05142

15

HRIFA025353a
GLYCINE-RICH CELL WALL STRUCTURAL PROTEIN 2 PRECURSOR.
1.0e-11:75:46
ORYZA SATIVA (RICE).
P29834

20

HRIFA025479a
PROTEASE DEGS PRECURSOR (EC 3.4.21.-).
3.0e-05:112:33
ESCHERICHIA COLI.
P31137

30

HRIFA025488a
PROCOLLAGEN ALPHA 1(III) CHAIN PRECURSOR (FRAGMENTS).
9.5e-05:104:40
MUS MUSCULUS (MOUSE).
P08121

35

HRIFA025492a
SERINE/THREONINE-SPECIFIC PROTEIN KINASE MINIBRAIN HOMOLOG (EC 2.7.1.-) (HP86) (DYRK).
1.8e-53:159:69
HOMO SAPIENS (HUMAN).
Q13627

40

HRIFA025636a
MITOCHONDRIAL RESPIRATORY CHAIN COMPLEXES ASSEMBLY PROTEIN RCA1 (EC 3.4.24.-) (TAT-BIND-
ING HOMOLOG 12).
4.7e-32:81:66
SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
P40341

45

HRIFA025695a
PEREGRIN (BR140 PROTEIN).
3.1e-40:227:43
HOMO SAPIENS (HUMAN).
P55201

55

HRIFA025703a

EP 1 130 094 A2

CELL SURFACE ANTIGEN 114/A10 PRECURSOR.
1.8e-08:71:42
MUS MUSCULUS (MOUSE).
P19467

5

HRIFA025706a
GLYCOSYLTRANSFERASE ALG2 (EC 2.4.1.-).
1.2e-28:111:44
SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
P43636

10

HRIFA025766a
CYTOCHROME B5.
4.2e-13:133:33
ORYCTOLAGUS CUNICULUS (RABBIT).
P00169

15

HRIFA025771a
HYPOTHETICAL 22.2 KD PROTEIN IN NSR1-TIF4631 INTERGENIC REGION.
6.7e-10:129:31
SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
P53288

20

HRIFA025778a
A-AGGLUTININ ATTACHMENT SUBUNIT PRECURSOR.
1.5e-05:212:30
SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
P32323

25

HRIFA025800a
HYPOTHETICAL 37.4 KD PROTEIN IN IRR1-TIM44 INTERGENIC REGION.
3.7e-18:165:33
SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
P40544

30

HRIFA025904a
COMPLEMENT RECEPTOR TYPE 1 PRECURSOR (C3B/C4B RECEPTOR) (CD35 ANTIGEN).
2.6e-05:211:28
HOMO SAPIENS (HUMAN).
P17927

35

HRIFA025907a
INTERFERON GAMMA UP-REGULATED I-5111 PROTEIN PRECURSOR (IGUPI-5111).
2.1e-38:176:38
HOMO SAPIENS (HUMAN).
Q06323

40

HRIFA025913a
DOLICHYL-PHOSPHATE-MANNOSE--PROTEIN MANNOSYLTRANSFERASE 4 (EC 2.4.1.109).
2.5e-32:185:37
SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
P46971

50

HRIFA025936a
TRANSCRIPTIONAL ACTIVATOR FE65.
3.4e-09:43:46
RATTUS NORVEGICUS (RAT).
P46933

55

EP 1 130 094 A2

HRIFA025966a
 SYNAPTOTAGMIN III.
 4.5e-05:93:33
 RATTUS NORVEGICUS (RAT).
 5 P40748

HRIFA025978a
 "GLUTENIN, HIGH MOLECULAR WEIGHT SUBUNIT DX5 PRECURSOR."
 3.5e-06:224:28
 10 TRITICUM AESTIVUM (WHEAT).
 P10388

HRIFA026089a
 BUTYROPHILIN PRECURSOR (BT).
 1.1e-12:146:29
 15 BOS TAURUS (BOVINE).
 P18892

HRIFA026121a
 FAS ANTIGEN LIGAND (APOPTOSIS ANTIGEN LIGAND) (APTL).
 9.7e-06:72:43
 20 HOMO SAPIENS (HUMAN).
 P48023

HRIFA026242a
 HYPOTHETICAL 73.0 KD PROTEIN IN CLA4-MID1 INTERGENIC REGION.
 7.4e-09:188:26
 25 SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
 P48566
 30

HRIFA026265a
 DNA BINDING PROTEIN S1FA.
 0.67:43:37
 ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).
 35 P42551

HRIFA026303a
 SALIVARY PROLINE-RICH PROTEIN PO PRECURSOR (ALLELE S).
 0.014:88:32
 40 HOMO SAPIENS (HUMAN).
 P10163

HRIFA026316a
 EBNA-2 NUCLEAR PROTEIN.
 1.5e-07:82:35
 45 EPSTEIN-BARR VIRUS (STRAIN B95-8) (HUMAN HERPESVIRUS 4).
 P12978

HRIFA026351a
 FLI-1 ONCOGENE (ERGB TRANSCRIPTION FACTOR).
 0.019:89:31
 50 HOMO SAPIENS (HUMAN).
 Q01543

HRIFA026364a
 PROBABLE G PROTEIN-COUPLED RECEPTOR EDG-1.
 8.3e-40:167:49
 55 RATTUS NORVEGICUS (RAT).

EP 1 130 094 A2

P48303

HRIFA026382a
T-CELL RECEPTOR BETA CHAIN PRECURSOR (ANA 11).
5 6.2e-10:135:38
ORYCTOLAGUS CUNICULUS (RABBIT).
P06333

HRIFA026465a
10 COLLAGEN ALPHA 1(IX) CHAIN PRECURSOR (FRAGMENTS).
8.6e-07:158:35
GALLUS GALLUS (CHICKEN).
P12106

HRIFA026496a
15 ZINC FINGER PROTEIN 140.
5.9e-24:122:52
HOMO SAPIENS (HUMAN).
P52738

HRIFA026519a
20 PROLINE-RICH PROTEIN MP-2 PRECURSOR.
1.3e-08:130:36
MUS MUSCULUS (MOUSE).
25 P05142

HRIFA026564a
GLYCOPROTEIN X PRECURSOR
1.8e-10:225:25
30 EQUINE HERPESVIRUS TYPE 1 (STRAIN AB4P) (EHV-1).
P28968

HRIFA026576a
35 *ADP,ATP CARRIER PROTEIN, HEART/SKELETAL MUSCLE ISOFORM T1 (ADP/ATP TRANSLOCASE 1) (AD-
ENINE NUCLEOTIDE TRANSLOCATOR 1) (ANT 1)."
1.7e-09:116:34
HOMO SAPIENS (HUMAN).
P12235

HRIFA026615a
40 REGULATORY PROTEIN E2.
0.0024:132:31
HUMAN PAPILLOMAVIRUS TYPE 9.
P36780

HRIFA026618a
45 PROTEIN Q300.
1.2e-05:27:66
MUS MUSCULUS (MOUSE).
50 Q02722

HRIFA026659a
SERINE/THREONINE-PROTEIN KINASE SGK (EC 2.7.1.-) (SERUM/GLUCOCORTICOID-REGULATED KI-
55 NASE).
2.0e-10:81:45
RATTUS NORVEGICUS (RAT).
Q06226

EP 1 130 094 A2

5	<p>HRIFA026764a MRC OX-45 SURFACE ANTIGEN PRECURSOR (BCM1 SURFACE ANTIGEN) (BLAST-1) (CD48). 3.4e-05:162:25 RATTUS NORVEGICUS (RAT). P10252</p>
10	<p>HRIFA026789a PUTATIVE GENERAL NEGATIVE REGULATOR OF TRANSCRIPTION C16C9.04C. 8.1e-22:175:38 SCHIZOSACCHAROMYCES POMBE (FISSION YEAST). Q09818</p>
15	<p>HRIFA026813a "PROTEIN KINASE C, MU TYPE (EC 2.7.1.-) (NPKC-MU)." 7.1e-89:256:67 HOMO SAPIENS (HUMAN). Q15139</p>
20	<p>HRIFA026860a MONOCARBOXYLATE TRANSPORTER 2 (MCT 2). 2.6e-19:103:43 MESOCRICETUS AURATUS (GOLDEN HAMSTER). P53988</p>
25	<p>HRIFA026923a CCAAT DISPLACEMENT PROTEIN (HOMEBOX PROTEIN CLOX) (CLOX-1) (FRAGMENT). 0.18:119:36 CANIS FAMILIARIS (DOG). P39881</p>
30	<p>HRIFA027012a "MANNOSYL-OLIGOSACCHARIDE ALPHA-1,2-MANNOSIDASE (EC 3.2.1.113) (MAN(9)-ALPHA-MANNOSIDASE) (ALPHA-MANNOSIDASE 1A)." 1.8e-44:234:41 MUS MUSCULUS (MOUSE). P45700</p>
40	<p>HRIFA027045a HYPOTHETICAL PROTEIN HI0519. 2.7e-27:181:38 HAEMOPHILUS INFLUENZAE. P44742</p>
45	<p>HRIFA027125a ZINC FINGER PROTEIN 133. 3.9e-33:70:61 HOMO SAPIENS (HUMAN). P52736</p>
50	<p>HRIFA027173a TRANS-ACTING TRANSCRIPTIONAL PROTEIN ICP0 (VMW118 PROTEIN). 0.15:137:27 HERPES SIMPLEX VIRUS (TYPE 2 / STRAIN HG52). P28284</p>
55	<p>HRIFA027179a EXCISION REPAIR PROTEIN ERCC-6 (COCKAYNE SYNDROME PROTEIN CSB). 3.6e-30:90:77</p>

EP 1 130 094 A2

- HOMO SAPIENS (HUMAN).
Q03468
- 5 HRIFA027187a
B-CELL GROWTH FACTOR PRECURSOR (BCGF-12 KD).
4.7e-11:44:61
HOMO SAPIENS (HUMAN).
P20931
- 10 HRIFA027327a
COLLAGEN ALPHA 1(X) CHAIN PRECURSOR.
3.8e-07:184:35
HOMO SAPIENS (HUMAN).
Q03692
- 15 HRIFA027329a
SALIVARY GLUE PROTEIN SGS-3 PRECURSOR.
9.1e-08:195:29
DROSOPHILA ERECTA (FRUIT FLY).
20 P13730
- HRIFA027355a
B-CELL GROWTH FACTOR PRECURSOR (BCGF-12 KD).
1.9e-06:33:72
25 HOMO SAPIENS (HUMAN).
P20931
- HRIFA027485a
COLLAGEN ALPHA 1(XI) CHAIN PRECURSOR.
30 0.00099:174:36
HOMO SAPIENS (HUMAN).
P12107
- HRIFA027536a
35 VITELLINE MEMBRANE PROTEIN VM26AB PRECURSOR (PROTEIN TU-4) (PROTEIN SV23).
0.0042:104:35
DROSOPHILA MELANOGASTER (FRUIT FLY).
P13238
- 40 HRIFA027549a
D(4) DOPAMINE RECEPTOR (D(2C) DOPAMINE RECEPTOR).
0.00023:101:44
HOMO SAPIENS (HUMAN).
P21917
- 45 HRIFA027622a
GUANOSINE-DIPHOSPHATASE (EC 3.6.1.42) (GDPASE).
2.2e-23:146:45
SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
50 P32621
- HRIFA027625a
CALCIUM-TRANSPORTING ATPASE 1 (EC 3.6.1.38) (GOLGI CA2+-ATPASE).
1.1e-57:220:54
55 SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
P13586
- HRIFA027644a

EP 1 130 094 A2

COLLAGEN ALPHA 1(I) CHAIN (FRAGMENTS).
 7.5e-05:72:40
 RATTUS NORVEGICUS (RAT).
 P02454

5

HRIFA027656a
 NON-RECEPTOR TYROSINE KINASE SPORE LYSIS A (EC 2.7.1.112) (TYROSINE-PROTEIN KINASE 1).
 1.6e-13:149:34
 DICTYOSTELIUM DISCOIDEUM (SLIME MOLD).
 P18160

10

HRIFA027673a
 CONNECTIVE TISSUE GROWTH FACTOR PRECURSOR.
 6.4e-06:47:57
 HOMO SAPIENS (HUMAN).
 P29279

15

HRIFA027681a
 SPORULATION-SPECIFIC PROTEIN 1 (EC 2.7.1.-).
 1.1e-13:158:31
 SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
 P08458

20

HRIFA027714a
 HYPOTHETICAL 146.8 KD PROTEIN C34E10.5 IN CHROMOSOME III.
 7.2e-06:146:30
 CAENORHABDITIS ELEGANS.
 P46580

25

HRIFA027722a
 SIGNAL RECOGNITION PARTICLE 68 KD PROTEIN (SRP68).
 2.7e-105:242:85
 CANIS FAMILIARIS (DOG).
 Q00004

30

HRIFA027860a
 SKIN SECRETORY PROTEIN XP2 PRECURSOR (APEG PROTEIN).
 8.3e-08:168:32
 XENOPUS LAEVIS (AFRICAN CLAWED FROG).
 P17437

35

HRIFA027867a
 STANNIOCALCIN PRECURSOR (STC) (CORPUSCLES OF STANNIUS PROTEIN) (CS) (HYPOCALCIN) (TEL-
 EOCALCIN).
 1.0:100:27
 ANGUILLA AUSTRALIS (AUSTRALIAN EEL).
 P18301

40

HRIFA027940a
 INHIBIN BETA C CHAIN PRECURSOR (ACTIVIN BETA-C CHAIN).
 8.7e-15:149:38
 HOMO SAPIENS (HUMAN).
 P55103

45

HRIFA028061a
 PUTATIVE SERINE/THREONINE-PROTEIN KINASE PKWA (EC 2.7.1.-).
 9.7e-07:157:26
 THERMOMONOSPORA CURVATA.

50

55

EP 1 130 094 A2

P49695

HRIFA028157a
HIGH-AFFINITY CATIONIC AMINO ACID TRANSPORTER-1 (CAT-1) (CAT1) (SYSTEM Y+ BASIC AMINO ACID
5 TRANSPORTER) (ECOTROPIC RETROVIRAL LEUKEMIA RECEPTOR HOMOLOG) (ERR) (ECOTROPIC RET-
ROVIRUS RECEPTOR HOMOLOG).
2.8e-71:201:68
HOMO SAPIENS (HUMAN).
P30825

10 HRIFA028187a
EBNA-1 NUCLEAR PROTEIN.
1.5e-09:131:38
EPSTEIN-BARR VIRUS (STRAIN B95-8) (HUMAN HERPESVIRUS 4).
15 P03211

HRIFA028262a
CELLULAR NUCLEIC ACID BINDING PROTEIN (CNBP).
7.2e-09:99:33
20 MUS MUSCULUS (MOUSE).
P53996

HRIFA028371a
PLASMINOGEN (EC 3.4.21.7) (FRAGMENT).
25 1.0e-08:103:33
RATTUS NORVEGICUS (RAT).
Q01177

HRIFA028402a
30 PUTATIVE SERINE/THREONINE-PROTEIN KINASE PKWA (EC 2.7.1.-).
3.2e-33:204:39
THERMOMONOSPORA CURVATA.
P49695

35 HRIFA028440a
COLLAGEN ALPHA 4(IV) CHAIN PRECURSOR.
1.9e-07:192:36
HOMO SAPIENS (HUMAN).
P53420

40 HRIFA028468a
CALCIUM/CALMODULIN-DEPENDENT PROTEIN KINASE TYPE IV CATALYTIC CHAIN (EC 2.7.1.123) (CAM
KINASE-GR) (CAMK IV) [CONTAINS: CALSPERMIN].
5.8e-32:178:44
45 RATTUS NORVEGICUS (RAT).
P13234

HRIFA028501a
50 VOLTAGE-GATED POTASSIUM CHANNEL PROTEIN KV1.6 (RCK2) (KV2).
6.3e-05:161:31
RATTUS NORVEGICUS (RAT).
P17659

55 HRIFA028511a
ANKYRIN HOMOLOG PRECURSOR.
3.0e-19:176:34
CHROMATIUM VINOSUM.
Q06527

EP 1 130 094 A2

5	<p>HRIFA028576a ACROSIN PRECURSOR (EC 3.4.21.10). 4.8e-08:78:46 ORYCTOLAGUS CUNICULUS (RABBIT). P48038</p>
10	<p>HRIFA028614a HISTIDINE-RICH GLYCOPROTEIN PRECURSOR. 1.0e-08:82:39 PLASMODIUM LOPHURAE. P04929</p>
15	<p>HRIFA028651a BAND 3 ANION TRANSPORT PROTEIN. 1.3e-18:156:32 GALLUS GALLUS (CHICKEN). P15575</p>
20	<p>HRIFA028790a PROBABLE CALCIUM-TRANSPORTING ATPASE 6 (EC 3.6.1.38). 5.0e-18:212:29 SACCHAROMYCES CEREVISIAE (BAKER'S YEAST). P39986</p>
25	<p>HRIFA028804a CCAAT-BINDING FACTOR (CBF). 0.98:232:23 MUS MUSCULUS (MOUSE). P53569</p>
30	<p>HRIFA028867a REGULATORY PROTEIN E2. 0.0057:124:31 HUMAN PAPILLOMAVIRUS TYPE 25. P36787</p>
35	<p>HRIFA028911a HYPOTHETICAL 118.4 KD PROTEIN IN BAT2-DAL5 INTERGENIC REGION PRECURSOR. 1.2e-09:206:33 SACCHAROMYCES CEREVISIAE (BAKER'S YEAST). P47179</p>
40	<p>HRIFA028983a HOMEOTIC GENE REGULATOR (BRAHMA PROTEIN). 0.0051:115:33 DROSOPHILA MELANOGASTER (FRUIT FLY). P25439</p>
45	<p>HRIFA029002a FIBRINOGEN BETA CHAIN. 3.2e-25:121:45 BOS TAURUS (BOVINE). P02676</p>
50	<p>HRIFA029050a RETINAL-CADHERIN PRECURSOR (R-CADHERIN) (R-CAD). 1.2e-10:134:32 GALLUS GALLUS (CHICKEN).</p>
55	

EP 1 130 094 A2

P24503

HRIFA029208a
RENAL TRANSCRIPTION FACTOR KID-1 (TRANSCRIPTION FACTOR 17).
1.4e-14:64:59
RATTUS NORVEGICUS (RAT).
Q02975

HRIFA029209a
"ALPHA-MANNOSIDASE II (EC 3.2.1.114) (MANNOSYL-OLIGOSACCHARIDE 1,3-1,6-ALPHA-MANNOSI-
DASE) (MAN II) (GOLGI ALPHA-MANNOSIDASE II)."
2.3e-12:114:37
MUS MUSCULUS (MOUSE).
P27046

HRIFA029256a
GAP JUNCTION BETA-2 PROTEIN (CONNEXIN 26) (CX26).
1.8e-35:89:75
HOMO SAPIENS (HUMAN).
P29033

HRIFA029263a
SARCALUMENIN PRECURSOR.
2.1e-16:161:31
ORYCTOLAGUS CUNICULUS (RABBIT).
P13666

HRIFA029278a
"SALIVARY PROLINE-RICH PROTEIN PRECURSOR (CLONES CP3, CP4 AND CP5) [CONTAINS: BASIC PEP-
TIDE IB-6" PEPTIDE P-H].
3.5e-10:204:32
HOMO SAPIENS (HUMAN).
P04280

HRIFA029285a
GLYCOPROTEIN 25L PRECURSOR (GP25L).
4.9e-58:197:55
CANIS FAMILIARIS (DOG).
P27869

HRIFA029317a
HIGH AFFINITY SULPHATE TRANSPORTER 2.
2.3e-25:83:50
STYLOSANTHES HAMATA.
P53392

HRIFA029327a
MITOCHONDRIAL 2-OXOGLUTARATE/MALATE CARRIER PROTEIN (OGCP).
9.1e-34:227:37
BOS TAURUS (BOVINE).
P22292

HRIFA029349a
CUTICLE COLLAGEN 12 PRECURSOR.
5.1e-09:190:33
CAENORHABDITIS ELEGANS.
P20630

EP 1 130 094 A2

HRIFA029393a
 PROBABLE G PROTEIN-COUPLED RECEPTOR APJ.
 9.7e-69:165:84
 HOMO SAPIENS (HUMAN).
 5 P35414

HRIFA029398a
 TRANS-ACTING TRANSCRIPTIONAL PROTEIN ICP0 (P135 PROTEIN) (IER 2.9/ER2.6).
 0.011:170:34
 10 BOVINE HERPESVIRUS TYPE 1 (STRAIN JURA).
 P29128

HRIFA029425a
 ALPHA CRYSTALLIN B CHAIN (ALPHA(B)-CRYSTALLIN).
 15 2.0e-08:99:32
 BOS TAURUS (BOVINE).
 P02510

HRIFA029434a
 20 "SALIVARY PROLINE-RICH PROTEIN PRECURSOR (CLONES CP3, CP4 AND CPS) [CONTAINS: BASIC PEP-
 TIDE IB-6" PEPTIDE P-H].
 2.6e-05:232:32
 HOMO SAPIENS (HUMAN).
 P04280
 25

HRIFA029440a
 SKIN SECRETORY PROTEIN XP2 PRECURSOR (APEG PROTEIN).
 0.00046:131:33
 XENOPUS LAEVIS (AFRICAN CLAWED FROG).
 30 P17437

HRIFA029460a
 SYNAPTOTAGMIN III.
 1.5e-08:102:35
 35 RATTUS NORVEGICUS (RAT).
 P40748

HRIFA029467a
 GLYCOPROTEIN X PRECURSOR.
 40 5.2e-07:182:31
 EQUINE HERPESVIRUS TYPE 1 (STRAIN AB4P) (EHV-1).
 P28968

HRIFA029508a
 45 PROPERDIN PRECURSOR.
 1.9e-06:218:32
 HOMO SAPIENS (HUMAN).
 P27918

HRIFA029511a
 50 POTASSIUM CHANNEL PROTEIN EAG.
 2.3e-66:139:61
 DROSOPHILA MELANOGASTER (FRUIT FLY).
 Q02280
 55

HRIFA029602a
 SEX-DETERMINING REGION Y PROTEIN (TESTIS-DETERMINING FACTOR).
 1.0:37:37

EP 1 130 094 A2

- SUS SCROFA (PIG).
P36393
- 5 HRIFA029649a
TRANS-ACTING TRANSCRIPTIONAL PROTEIN ICP0 (VMW118 PROTEIN).
0.30:99:34
HERPES SIMPLEX VIRUS (TYPE 2 / STRAIN HG52).
P28284
- 10 HRIFA029715a
GROWTH ARREST AND DNA-DAMAGE-INDUCIBLE PROTEIN GADD153 (DNA-DAMAGE INDUCIBLE TRAN-
SCRIPT 3) (DDIT3) (C/EBP-HOMOLOGOUS PROTEIN) (CHOP).
0.54:95:30
HOMO SAPIENS (HUMAN).
15 P35638
- HRIFA029730a
HISTIDINE-RICH GLYCOPROTEIN PRECURSOR.
3.8e-05:131:29
20 PLASMODIUM LOPHURAE.
P04929
- HRIFA029792a
PUTATIVE SERINE/THREONINE-PROTEIN KINASE PKWA (EC 2.7.1.-).
25 9.0e-09:178:30
THERMOMONOSPORA CURVATA.
P49695
- HRIFA029802a
30 TRAM PROTEIN (TRANSLOCATING CHAIN-ASSOCIATING MEMBRANE PROTEIN).
7.2e-73:204:69
CANIS FAMILIARIS (DOG).
Q01685
- 35 HRIFA029866a
PROTEIN KINASE BYR2 (EC 2.7.1.-) (PROTEIN KINASE STE8) (MAPK KINASE KINASE) (MAPKKK).
1.2e-27:144:45
SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).
P28829
40
- HRIFA029932a
F-SPONDIN PRECURSOR.
9.1e-24:191:37
XENOPUS LAEVIS (AFRICAN CLAWED FROG).
45 P35447
- HRIFA030025a
ENDOSOMAL P24A PROTEIN PRECURSOR (70 KD ENOMEMBRANE PROTEIN) (PHEROMONE ALPHA-
FACTOR TRANSPORTER) (ACIDIC 24 KD LATE ENDOCYTIC INTERMEDIATE COMPONENT).
50 1.0e-11:138:31
SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
P32802
- 55 HRIFA030045a
SARCALUMENIN PRECURSOR.
2.4e-20:151:32
ORYCTOLAGUS CUNICULUS (RABBIT).
P13666

EP 1 130 094 A2

HRIFA030103a
HYPOTHETICAL 57.5 KD PROTEIN IN VMA7-RPS25A INTERGENIC REGION.
2.1e-05:215:29
SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
5 P53214

HRIFA030106a
SCO-SPONDIN (FRAGMENT).
0.53:60:36
10 BOS TAURUS (BOVINE).
P98167

HRIFA030147a
PUTATIVE MITOCHONDRIAL CARRIER YGR096W.
15 1.8e-10:93:34
SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
P53257

HRIFA030250a
20 ENAMELIN (TUFTELIN).
3.7e-108:250:86
BOS TAURUS (BOVINE).
P27628

HRIFA030264a
25 SODIUM/GLUCOSE COTRANSPORTER 1 (NA(+)/GLUCOSE COTRANSPORTER 1) (HIGH AFFINITY SODI-
UM-GLUCOSE COTRANSPORTER).
3.3e-09:119:27
ORYCTOLAGUS CUNICULUS (RABBIT).
30 P11170

HRIFA030342a
ESTRADIOL 17 BETA-DEHYDROGENASE 3 (EC 1.1.1.62) (17-BETA-HSD 3) (TESTICULAR 17-BETA-HY-
DROXYSTEROID DEHYDROGENASE).
35 1.5e-42:203:49
HOMO SAPIENS (HUMAN).
P37058

HRIFA030370a
40 HYPOTHETICAL 56.2 KD PROTEIN CY31.25C.
8.0e-12:88:48
MYCOBACTERIUM TUBERCULOSIS.
Q10555

HRIFA030371a
45 "PROTEIN KINASE C, MU TYPE (EC 2.7.1.-) (NPKC-MU)."
1.6e-68:228:59
HOMO SAPIENS (HUMAN).
Q15139
50

HRIFA030381a
COLLAGEN 1(X) CHAIN PRECURSOR.
3.0e-05:204:30
GALLUS GALLUS (CHICKEN).
55 P08125

HRIFA030385a
COLLAGEN ALPHA 1(X) CHAIN PRECURSOR.

EP 1 130 094 A2

0.029:162:31
HOMO SAPIENS (HUMAN).
Q03692

5 HRIFA030411a
SERINE PALMITOYLTRANSFERASE 2 (EC 2.3.1.50) (LONG CHAIN BASE BIOSYNTHESIS PROTEIN 2) (SPT
2).
1.2e-27:115:53
SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).
10 Q09925

HRIFA030448a
PUTATIVE SERINE/THREONINE-PROTEIN KINASE P78 (EC 2.7.1.-).
2.5e-92:225:77
15 HOMO SAPIENS (HUMAN).
P27448

HRIFA030456a
PROLINE-RICH PROTEIN MP-2 PRECURSOR.
9.3e-08:127:35
20 MUS MUSCULUS (MOUSE).
P05142

HRIFA030461a
CUTICLE COLLAGEN 12 PRECURSOR.
0.046:140:31
25 CAENORHABDITIS ELEGANS.
P20630

HRIFA030472a
NUC-1 NEGATIVE REGULATORY PROTEIN PREG.
0.0030:98:31
30 NEUROSPORA CRASSA.
Q06712

35 HRIFA030509a
"INTERFERON-INDUCED, DOUBLE-STRANDED RNA-ACTIVATED PROTEIN KINASE (EC 2.7.1.-) (INTER-
FERON-INDUCIBLE RNA-DEPENDENT PROTEIN KINASE) (P68 KINASE) (P1/EIF-2A PROTEIN KINASE)."
2.5e-09:65:43
40 HOMO SAPIENS (HUMAN).
P19525

HRIFA030511a
T-LYMPHOCYTE MATURATION-ASSOCIATED PROTEIN.
0.00010:99:33
45 HOMO SAPIENS (HUMAN).
P21145

HRIFA030545a
PROBABLE SERINE/THREONINE-PROTEIN KINASE YNL020C (EC 2.7.1.-).
7.6e-21:165:35
50 SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
P53974

55 HRIFA030566a
"GLUCOAMYLASE S1/S2 PRECURSOR (EC 3.2.1.3) (GLUCAN 1,4-ALPHA-GLUCOSIDASE) (1,4-ALPHA-D-
GLUCAN GLUCOHYDROLASE)."
2.7e-07:221:30

EP 1 130 094 A2

SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
P08640

5 HRIFA030599a
GLYCOPROTEIN X PRECURSOR.
2.8e-05:236:27
EQUINE HERPESVIRUS TYPE 1 (STRAIN AB4P) (EHV-1).
P28968

10 HRIFA030629a
PROTEIN DISULFIDE ISOMERASE PRECURSOR (PDI) (EC 5.3.4.1) (PROLYL 4-HYDROXYLASE BETA SUB-
UNIT) (CELLULAR THYROID HORMONE BINDING PROTEIN) (P55).
3.5e-16:115:38
BOS TAURUS (BOVINE).
15 P05307

HRIFA030642a
SULFATED SURFACE GLYCOPROTEIN 185 (SSG 185).
2.5e-12:93:47
20 VOLVOX CARTERI.
P21997

HRIFA030662a
NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 1 (EC 1.6.5.3).
25 9.1e-120:279:83
HOMO SAPIENS (HUMAN).
P03886

HRIFA030839a
30 HYPOTHETICAL GENE 51 MEMBRANE PROTEIN.
1.0:66:27
ICTALURID HERPESVIRUS 1 (CHANNEL CATFISH VIRUS) (CCV).
Q00135

35 HRIFA031091a
PROTEIN Q300.
0.0042:27:62
MUS MUSCULUS (MOUSE).
Q02722

40 HRIFA031126a
P2Y PURINOCEPTOR 5 (P2Y5) (PURINERGIC RECEPTOR 5) (RB INTRON ENCODED G-PROTEIN COUPLED
RECEPTOR).
1.3e-06:70:34
45 HOMO SAPIENS (HUMAN).
P43657

HRIFA031249a
ACIDIC PROLINE-RICH PROTEIN PRECURSOR (CLONE PRP33).
50 5.9e-05:166:31
RATTUS NORVEGICUS (RAT).
P04474

HRIFA031336a
55 CCAAT-BINDING TRANSCRIPTION FACTOR SUBUNIT A (CBF-A) (NF-Y PROTEIN CHAIN B) (NF-YB) (CAAT-
BOX DNA BINDING PROTEIN SUBUNIT B).
6.6e-15:97:38
PETROMYZON MARINUS (SEA LAMPREY).

EP 1 130 094 A2

P25210

HRIFA031395a
COLD SHOCK PROTEIN CSPB (FRAGMENT).
5 0.95:32:40
BACILLUS GLOBISPORUS.
P41018

HRIFA031397a
10 REGULATORY PROTEIN E2.
0.0077:145:35
HUMAN PAPILLOMAVIRUS TYPE 47.
P22420

HRIFA031438a
15 GLUCOSE REPRESSION MEDIATOR PROTEIN.
1.3e-06:176:26
SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
P14922

HRIFA031869a
20 TRANSCRIPTION FACTOR HES-1 (HAIRY AND ENHANCER OF SPLIT 1) (HAIRY-LIKE) (RHL).
1.7e-18:163:41
RATTUS NORVEGICUS (RAT).
25 Q04666

HRIFA031935a
EXTENSIN PRECURSOR (PROLINE-RICH GLYCOPROTEIN).
1.8e-06:192:32
30 ZEA MAYS (MAIZE).
P14918

HRIFA031986a
35 SERINE/THREONINE-PROTEIN KINASE PAK-ALPHA (EC 2.7.1.-) (P68-PAK) (P21-ACTIVATED KINASE) (ALPHA-PAK) (PROTEIN KINASE MUK2).
2.4e-49:222:47
RATTUS NORVEGICUS (RAT).
P35465

HRIFA032009a
40 PROBABLE G PROTEIN-COUPLED RECEPTOR FROM T-CELLS PRECURSOR (GLUCOCORTICOID-INDUCED RECEPTOR).
1.0e-17:118:36
MUS MUSCULUS (MOUSE).
45 P30731

HRIFA032011a
MUSCARINIC ACETYLCHOLINE RECEPTOR M4.
7.8e-35:184:32
50 HOMO SAPIENS (HUMAN).
P08173

HRIFA032070a
MITOCHONDRIAL RNA SPLICING PROTEIN MSR4.
55 2.1e-18:107:44
SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
P23500

EP 1 130 094 A2

HRIFA032073a
 SECRETAGRANIN III PRECURSOR (SGIII).
 9.7e-69:182:76
 MUS MUSCULUS (MOUSE).
 5 P47867

HRIFA032079a
 HYPOTHETICAL 49.3 KD PROTEIN C30D11.06C IN CHROMOSOME I.
 3.5e-12:96:39
 10 SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).
 Q09906

HRIFA032097a
 GLYCOPROTEIN J.
 15 0.023:61:32
 HERPES SIMPLEX VIRUS (TYPE 1 / STRAIN 17).
 P06480

HRIFA032161a
 20 CCAAT/ENHANCER BINDING PROTEIN DELTA (C/EBP DELTA) (NUCLEAR FACTOR NF-IL6-BETA) (NF-IL6-BETA).
 0.22:56:42
 HOMO SAPIENS (HUMAN).
 P49716
 25

HRIFA032186a
 D-BINDING PROTEIN (DBP) (ALBUMIN D BOX-BINDING PROTEIN) (TAXREB302).
 0.86:50:38
 HOMO SAPIENS (HUMAN).
 30 Q10586

HRIFA032224a
 HYPOTHETICAL 52.7 KD PROTEIN C38C10.2 IN CHROMOSOME III.
 2.6e-43:196:45
 35 CAENORHABDITIS ELEGANS.
 Q03567

HRIFA032257a
 GLUCOSE REPRESSION MEDIATOR PROTEIN.
 40 4.7e-07:204:25
 SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
 P14922

HRIFA032274a
 45 ZINC FINGER PROTEIN 38 (ZFP-38) (CTFIN51) (TRANSCRIPTION FACTOR RU49).
 7.8e-60:163:74
 MUS MUSCULUS (MOUSE).
 Q07231

HRIFA032275a
 50 CELL DIVISION CONTROL PROTEIN 28 (EC 2.7.1.-).
 7.2e-41:179:38
 SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
 P00546
 55

HRIFA032360a
 HYPOTHETICAL 84.3 KD PROTEIN ZK945.10 IN CHROMOSOME II.
 3.0e-05:198:28

EP 1 130 094 A2

CAENORHABDITIS ELEGANS.
Q09625

5 HRIFA032389a
EBNA-1 NUCLEAR PROTEIN.
1.3e-05:86:39
EPSTEIN-BARR VIRUS (STRAIN B95-8) (HUMAN HERPESVIRUS 4). P03211

10 HRIFA032433a
GONADOTROPIN-RELEASING HORMONE RECEPTOR (GNRH-R).
3.1e-14:54:53
RATTUS NORVEGICUS (RAT).
P30969

15 HRIFA032453a
BUTYROPHILIN PRECURSOR (BT).
5.9e-13:162:32
BOS TAURUS (BOVINE).
P18892

20 HRIFA032478a
GLYCOPROTEIN X PRECURSOR.
3.8e-06:253:28
EQUINE HERPESVIRUS TYPE 1 (STRAIN AB4P) (EHV-1).
25 P28968

HRIFA032506a
COLLAGEN ALPHA 3(VI) CHAIN PRECURSOR.
1.2e-06:226:34
30 HOMO SAPIENS (HUMAN).
P12111

HRIFA032511a
COLLAGEN ALPHA 1(XVI) CHAIN PRECURSOR.
35 8.7e-09:229:34
HOMO SAPIENS (HUMAN).
Q07092

HRIFA032530a
40 SKIN SECRETORY PROTEIN XP2 PRECURSOR (APEG PROTEIN).
9.0e-05:159:33
XENOPUS LAEVIS (AFRICAN CLAWED FROG).
P17437

45 HRIFA032587a
SYNAPTOTAGMIN (P65).
3.2e-08:72:52
APLYSIA CALIFORNICA (CALIFORNIA SEA HARE).
P41823

50 HRIFA032605a
ANTIGEN PEPTIDE TRANSPORTER 1 (APT1) (PEPTIDE TRANSPORTER TAP1) (PEPTIDE TRANSPORTER
PSF1) (PEPTIDE SUPPLY FACTOR 1) (PSF-1) (PEPTIDE TRANSPORTER INVOLVED IN ANTIGEN PROCESS-
55 8.4e-37:192:41
HOMO SAPIENS (HUMAN).
Q03518

EP 1 130 094 A2

HRIFA032642a
 PROLINE-RICH PROTEIN MP-2 PRECURSOR.
 5.0e-05:127:33
 MUS MUSCULUS (MOUSE).
 5 P05142

HRIFA032696a
 COLLAGEN ALPHA 1(II) CHAIN (FRAGMENTS).
 1.4e-13:200:38
 10 BOS TAURUS (BOVINE).
 P02459

HRIFA032730a
 K-GLYPICAN PRECURSOR.
 15 4.8e-67:180:68
 MUS MUSCULUS (MOUSE).
 P51655

HRIFA032820a
 20 GLUTAMIC ACID-RICH PROTEIN PRECURSOR.
 7.5e-05:192:23
 PLASMODIUM FALCIPARUM (ISOLATE FC27 / PAPUA NEW GUINEA).
 P13816

25 Homology search result 2

[0290] Homology of representative sequences of the 5'-end cluster to the data in Swiss-Prot database

[0291] Representative sequence of the 5'-end cluster exhibiting relatively high homology (304 cluster: "exhibiting relatively high homology" means that the P value is 10^{-10} or less)

30 HRIFA000327a, HRIFA000432a, HRIFA000553a, HRIFA000564a, HRIFA000631a, HRIFA000683a, HRIFA000776a,
 HRIFA000814a, HRIFA001132a, HRIFA001138a, HRIFA001337a, HRIFA001341a, HRIFA001489a, HRIFA001712a,
 HRIFA001720a, HRIFA001942a, HRIFA001975a, HRIFA001984a, HRIFA002384a, HRIFA002503a, HRIFA002743a,
 HRIFA002766a, HRIFA002805a, HRIFA002891a, HRIFA002919a, HRIFA002980a, HRIFA003063a, HRIFA003093a,
 HRIFA003635a, HRIFA004006a, HRIFA004034a, HRIFA004112a, HRIFA004426a, HRIFA004490a, HRIFA004523a,
 35 HRIFA004663a, HRIFA004696a, HRIFA004714a, HRIFA004745a, HRIFA004919a, HRIFA005184a, HRIFA005231a,
 HRIFA005240a, HRIFA005271a, HRIFA005372a, HRIFA005392a, HRIFA005409a, HRIFA005420a, HRIFA005438a,
 HRIFA005462a, HRIFA005644a, HRIFA005720a, HRIFA005732a, HRIFA005760a, HRIFA005781a, HRIFA006183a,
 HRIFA006494a, HRIFA006510a, HRIFA006566a, HRIFA006586a, HRIFA006596a, HRIFA006649a, HRIFA006667a,
 HRIFA006730a, HRIFA006926a, HRIFA007013a, HRIFA007219a, HRIFA007228a, HRIFA007243a, HRIFA007352a,
 40 HRIFA007424a, HRIFA007435a, HRIFA007463a, HRIFA007493a, HRIFA007571a, HRIFA007659a, HRIFA007722a,
 HRIFA007745a, HRIFA008000a, HRIFA008200a, HRIFA008284a, HRIFA008314a, HRIFA008362a, HRIFA008459a,
 HRIFA008483a, HRIFA008547a, HRIFA008611a, HRIFA008661a, HRIFA008717a, HRIFA008784a, HRIFA008981a,
 HRIFA009101a, HRIFA009171a, HRIFA009220a, HRIFA009451a, HRIFA009482a, HRIFA009783a, HRIFA009881a,
 HRIFA010085a, HRIFA010090a,
 45 HRIFA010130a, HRIFA010319a, HRIFA010394a, HRIFA010460a, HRIFA010790a, HRIFA010975a, HRIFA011016a,
 HRIFA011179a, HRIFA011197a, HRIFA011449a, HRIFA011659a, HRIFA011947a, HRIFA012278a, HRIFA012584a,
 HRIFA012625a, HRIFA012692a, HRIFA012795a, HRIFA012885a, HRIFA012914a, HRIFA012969a, HRIFA012990a,
 HRIFA013254a, HRIFA013265a, HRIFA013276a, HRIFA013376a, HRIFA013477a, HRIFA013586a, HRIFA013726a,
 HRIFA013744a, HRIFA013911a, HRIFA014006a, HRIFA014185a, HRIFA014336a, HRIFA014465a, HRIFA014500a,
 50 HRIFA014561a, HRIFA014568a, HRIFA014621a, HRIFA014688a, HRIFA014819a, HRIFA014951a, HRIFA014967a,
 HRIFA015063a, HRIFA015070a, HRIFA015246a, HRIFA015423a, HRIFA015453a, HRIFA015486a, HRIFA015506a,
 HRIFA015536a, HRIFA015547a, HRIFA015568a, HRIFA015756a, HRIFA015811a, HRIFA016070a, HRIFA016290a,
 HRIFA016430a, HRIFA016654a, HRIFA016758a, HRIFA017031a, HRIFA017257a, HRIFA017295a, HRIFA017312a,
 HRIFA017703a, HRIFA017855a, HRIFA018092a, HRIFA018131a, HRIFA018134a, HRIFA018580a, HRIFA018827a,
 55 HRIFA018904a, HRIFA018993a, HRIFA019105a, HRIFA019136a, HRIFA019175a, HRIFA019262a, HRIFA019466a,
 HRIFA019867a, HRIFA019869a, HRIFA020272a, HRIFA020335a, HRIFA020349a, HRIFA020862a, HRIFA021213a,
 HRIFA021398a, HRIFA021499a, HRIFA021637a, HRIFA021651a, HRIFA021754a, HRIFA021781a, HRIFA022065a,
 HRIFA022139a, HRIFA022166a, HRIFA022177a, HRIFA022182a, HRIFA022227a, HRIFA022249a, HRIFA022265a,

EP 1 130 094 A2

HRIFA022328a, HRIFA022423a, HRIFA022528a, HRIFA022546a, HRIFA022564a, HRIFA022616a, HRIFA022671a, HRIFA022691a, HRIFA022707a, HRIFA022729a, HRIFA022737a, HRIFA022776a, HRIFA022875a, HRIFA022895a, HRIFA023007a, HRIFA023227a, HRIFA023257a, HRIFA023304a, HRIFA023464a, HRIFA023767a, HRIFA023923a, HRIFA024132a, HRIFA024255a, HRIFA024392a, HRIFA024423a, HRIFA024504a, HRIFA024718a, HRIFA024767a, HRIFA024937a, HRIFA024994a, HRIFA025046a, HRIFA025250a, HRIFA025261a, HRIFA025353a, HRIFA025492a, HRIFA025636a, HRIFA025695a, HRIFA025706a, HRIFA025766a, HRIFA025800a, HRIFA025907a, HRIFA025913a, HRIFA026089a, HRIFA026364a, HRIFA026496a, HRIFA026789a, HRIFA026813a, HRIFA026860a, HRIFA027012a, HRIFA027045a, HRIFA027125a, HRIFA027179a, HRIFA027187a, HRIFA027622a, HRIFA027625a, HRIFA027656a, HRIFA027681a, HRIFA027722a, HRIFA027940a, HRIFA028157a, HRIFA028402a, HRIFA028468a, HRIFA028511a, HRIFA028651a, HRIFA028790a, HRIFA029002a, HRIFA029208a, HRIFA029209a, HRIFA029256a, HRIFA029263a, HRIFA029285a, HRIFA029317a, HRIFA029327a, HRIFA029393a, HRIFA029511a, HRIFA029802a, HRIFA029866a, HRIFA029932a, HRIFA030025a, HRIFA030045a, HRIFA030250a, HRIFA030342a, HRIFA030370a, HRIFA030371a, HRIFA030411a, HRIFA030448a, HRIFA030545a, HRIFA030629a, HRIFA030642a, HRIFA030662a, HRIFA031336a, HRIFA031869a, HRIFA031986a, HRIFA032009a, HRIFA032011a, HRIFA032070a, HRIFA032073a, HRIFA032079a, HRIFA032224a, HRIFA032274a, HRIFA032275a, HRIFA032433a, HRIFA032453a, HRIFA032605a, HRIFA032696a, HRIFA032730a,

Homology search result 3

[0292] Representative sequence of the 5'-end cluster exhibiting relatively low homology (221 cluster: "exhibiting relatively low homology" means that the P value is higher than 10^{-10} and 10^{-4} or less)
 HRIFA000016a, HRIFA000071a, HRIFA000116a, HRIFA000123a, HRIFA000264a, HRIFA000415a, HRIFA000446a, HRIFA000695a, HRIFA000845a, HRIFA001971a, HRIFA002063a, HRIFA002102a, HRIFA002284a, HRIFA002309a, HRIFA002694a, HRIFA002762a, HRIFA002787a, HRIFA003055a, HRIFA003340a, HRIFA003402a, HRIFA003504a, HRIFA003892a, HRIFA003946a, HRIFA004162a, HRIFA004401a, HRIFA004780a, HRIFA005072a, HRIFA005102a, HRIFA005214a, HRIFA005255a, HRIFA005300a, HRIFA005369a, HRIFA005702a, HRIFA005728a, HRIFA005944a, HRIFA006298a, HRIFA006448a, HRIFA006572a, HRIFA006633a, HRIFA006642a, HRIFA007068a, HRIFA007244a, HRIFA007262a, HRIFA007512a, HRIFA007532a, HRIFA007565a, HRIFA007728a, HRIFA007909a, HRIFA008174a, HRIFA008426a, HRIFA008596a, HRIFA008790a, HRIFA008989a, HRIFA009578a, HRIFA009825a, HRIFA009852a, HRIFA009983a, HRIFA010005a, HRIFA010078a, HRIFA010152a, HRIFA010301a, HRIFA010361a, HRIFA010425a, HRIFA010466a, HRIFA010799a, HRIFA011580a, HRIFA011820a, HRIFA012167a, HRIFA012354a, HRIFA012427a, HRIFA012436a, HRIFA012515a, HRIFA012702a, HRIFA012737a, HRIFA013135a, HRIFA013235a, HRIFA013279a, HRIFA013589a, HRIFA013620a, HRIFA013919a, HRIFA013932a, HRIFA014056a, HRIFA014111a, HRIFA014133a, HRIFA014396a, HRIFA014397a, HRIFA014598a, HRIFA014702a, HRIFA014868a, HRIFA015219a, HRIFA015995a, HRIFA016214a, HRIFA016240a, HRIFA016255a, HRIFA016639a, HRIFA016669a, HRIFA016963a, HRIFA017457a, HRIFA017643a, HRIFA017670a, HRIFA017801a, HRIFA017836a, HRIFA017921a, HRIFA018238a, HRIFA018262a, HRIFA018287a, HRIFA018666a, HRIFA018688a, HRIFA018754a, HRIFA018794a, HRIFA018870a, HRIFA018931a, HRIFA019412a, HRIFA019490a, HRIFA019498a, HRIFA019532a, HRIFA019651a, HRIFA0201440, HRIFA020184a, HRIFA020453a, HRIFA020693a, HRIFA020707a, HRIFA020748a, HRIFA021061a, HRIFA021224a, HRIFA021494a, HRIFA021794a, HRIFA021855a, HRIFA021906a, HRIFA022156a, HRIFA022203a, HRIFA022234a, HRIFA022702a, HRIFA022728a, HRIFA022782a, HRIFA022865a, HRIFA022890a, HRIFA022985a, HRIFA023048a, HRIFA023069a, HRIFA023129a, HRIFA023154a, HRIFA023212a, HRIFA023489a, HRIFA023634a, HRIFA023894a, HRIFA024088a, HRIFA024197a, HRIFA024218a, HRIFA024473a, HRIFA024482a, HRIFA024543a, HRIFA025327a, HRIFA025479a, HRIFA025488a, HRIFA025703a, HRIFA025771a, HRIFA025778a, HRIFA025904a, HRIFA025966a, HRIFA025978a, HRIFA026121a, HRIFA026242a, HRIFA026316a, HRIFA026382a, HRIFA026465a, HRIFA026519a, HRIFA026564a, HRIFA026576a, HRIFA026618a, HRIFA026659a, HRIFA026764a, HRIFA027327a, HRIFA027329a, HRIFA027355a, HRIFA027644a, HRIFA027673a, HRIFA027714a, HRIFA027860a, HRIFA028061a, HRIFA028187a, HRIFA028262a, HRIFA028371a, HRIFA028440a, HRIFA028501a, HRIFA028576a, HRIFA028614a, HRIFA028911a, HRIFA029050a, HRIFA029278a, HRIFA029349a, HRIFA029425a, HRIFA029434a, HRIFA029460a, HRIFA029467a, HRIFA029508a, HRIFA029730a, HRIFA029792a, HRIFA030103a, HRIFA030147a, HRIFA030264a, HRIFA030381a, HRIFA030456a, HRIFA030509a, HRIFA030511a, HRIFA030566a, HRIFA030599a, HRIFA031126a, HRIFA031249a, HRIFA031438a, HRIFA031935a, HRIFA032257a, HRIFA032360a, HRIFA032389a, HRIFA032478a, HRIFA032506a, HRIFA032511a, HRIFA032530a, HRIFA032587a, HRIFA032642a, HRIFA032820a,

Homology search result 4

[0293] Representative sequence of the 5'-end cluster exhibiting low homology (115 cluster: "exhibiting low homology"

EP 1 130 094 A2

means that the P value is higher than 10^{-4} and 1 or less)

HRIFA001099a, HRIFA001200a, HRIFA001413a, HRIFA001439a, HRIFA001558a, HRIFA001866a, HRIFA001972a,
 HRIFA002689a, HRIFA003357a, HRIFA003592a, HRIFA003640a, HRIFA003883a, HRIFA005296a, HRIFA005500a,
 HRIFA005540a, HRIFA006250a, HRIFA006609a, HRIFA006798a, HRIFA007032a, HRIFA007152a, HRIFA007547a,
 5 HRIFA007829a, HRIFA007985a, HRIFA008212a, HRIFA008252a, HRIFA008976a, HRIFA009071a, HRIFA009123a,
 HRIFA009136a, HRIFA009339a, HRIFA009762a, HRIFA010176a, HRIFA010490a, HRIFA010736a, HRIFA010859a,
 HRIFA010891a, HRIFA010988a, HRIFA011105a, HRIFA011128a, HRIFA011484a, HRIFA011512a, HRIFA011926a,
 HRIFA012069a, HRIFA012151a, HRIFA013092a, HRIFA013103a, HRIFA013980a, HRIFA014024a, HRIFA014590a,
 HRIFA014620a, HRIFA015122a, HRIFA015351a, HRIFA015802a, HRIFA015902a, HRIFA015947a, HRIFA016599a,
 10 HRIFA017146a, HRIFA017190a, HRIFA017456a, HRIFA017791a, HRIFA017818a, HRIFA018447a, HRIFA019437a,
 HRIFA019958a, HRIFA020883a, HRIFA021007a, HRIFA021040a, HRIFA021445a, HRIFA021543a, HRIFA021620a,
 HRIFA021787a, HRIFA022055a, HRIFA022335a, HRIFA022348a, HRIFA022411a, HRIFA022462a, HRIFA022493a,
 HRIFA022714a, HRIFA023434a, HRIFA024185a, HRIFA024305a, HRIFA024884a, HRIFA024893a, HRIFA024978a,
 HRIFA025033a, HRIFA025290a, HRIFA026265a, HRIFA026303a, HRIFA026351a, HRIFA026615a, HRIFA026923a,
 15 HRIFA027173a, HRIFA027485a, HRIFA027536a, HRIFA027549a, HRIFA027867a, HRIFA028804a, HRIFA028867a,
 HRIFA028983a, HRIFA029398a,
 HRIFA029440a, HRIFA029602a, HRIFA029649a, HRIFA029715a, HRIFA030106a, HRIFA030385a, HRIFA030461a,
 HRIFA030472a, HRIFA030839a, HRIFA031091a, HRIFA031395a, HRIFA031397a, HRIFA032097a, HRIFA032161a,
 HRIFA032186a,

20

Homology search result 5

[0294] The result of the homology search in the SwissProt using the clone sequences of the 5'-ends.

25 Indicated are from the top,
 the name of the clone sequence,
 definition of the top hit data,
 the P-value: the length of the sequence used for comparison (nucleotide):similarity (%),
 the organism of which the top hit data is obtained,
 30 the Accession No. of the top hit data.

[0295] Data were not shown for the clones in which the P-value was higher than 1.

35 F-BNGH41000020
 NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 1 (EC 1.6.5.3).
 1.2e-119:279:83
 HOMO SAPIENS (HUMAN).
 P03886

40 F-BNGH41000087
 PROPERDIN PRECURSOR.
 2.5e-06:218:32
 HOMO SAPIENS (HUMAN).
 P27918

45 F-BNGH41000091
 POTASSIUM CHANNEL PROTEIN EAG.
 3.1e-66:139:61
 DROSOPHILA MELANOGASTER (FRUIT FLY).
 50 Q02280

F-HEMBA1000006
 S-ANTIGEN PROTEIN PRECURSOR.
 3.0e-05:164:31
 55 PLASMODIUM FALCIPARUM (ISOLATE V1).
 P09593

F-HEMBA1000121

EP 1 130 094 A2

Homo sapiens DNA recombination and repair protein (MRE11B) mRNA, complete cds
0.00075:456:59
Hs.153855:AF022778

- 5 R-Y79AA1002213
Human mRNA for KIAA0392 gene, partial cds
6.2e-45:304:85
Hs.40100:AB002390
- 10 R-Y79AA1002334
ESTs
7.7e-91:495:92
Hs.90804:W28091
- 15 R-Y79AA1002373
Human kpni repeat mrna (cdna clone pcd-kpni-8), 3' end
5.2e-98:545:91
Hs.103948:K00627
- 20 R-Y79AA1002376
ESTs
2.0e-91:455:97
Hs.153375:AI287812
- 25 R-Y79AA1002378
ESTs, Highly similar to ZINC FINGER PROTEIN ZFP-35 [Mus musculus]
9.4e-15:131:83
Hs.20082:W89121
- 30 R-Y79AA1002381
ESTs, Highly similar to CELL DIVISION CONTROL PROTEIN 2 HOMOLOG [Plasmodium falciparum (isolate k1/thailand)]
1.5e-104:531:95
Hs.26322:AA156858

35 Homology search result 10

[0305] Data obtained by the homology search for full length nucleotide sequences and deduced amino acid sequences. In the result of the search shown below, both units, aa and bp, are used as length units for the sequences to be compared. Each data includes Clone name, Definition in matching data, P value, Length of sequence to be compared, Homology, and Accession number (No.) of matching data. These items are shown in this order, separated by a double-slash mark, //.

- 45 C-HEMBA1000006//Homo sapiens mRNA; cDNA DKFZp564G1762 (from clone DKFZp564G1762).//0//1230bp//92%//AB026894
C-nnnnnnnnnnnnn//GAMETOGENESIS EXPRESSED PROTEIN GEG-154.//2.30E-71//344aa//50%//P50636
C-HEMBA1000121//HYPOTHETICAL 68.7 KD PROTEIN ZK757.1 IN CHROMOSOME III.//4.80E-05//83aa//27%//P34679
C-HEMBA1000128//PATHOGENESIS-RELATED PROTEIN 1 PRECURSOR (PR-1).//3.20E-07//89aa//34%//P33154
- 50 C-HEMBA1000275
C-HEMBA1000300
C-HEMBA1000349//ATP-BINDING CASSETTE TRANSPORTER 1.//5.30E-65//352aa//39%//P41233
C-HEMBA1000443//Homo sapiens CGI-96 protein mRNA, complete cds.//4.70E-129//686bp//91%//AF151854
- 55 C-HEMBA1000590//Homo sapiens mRNA for matrilin-4, partial.//2.00E-273//1254bp//99%//AJ007581
C-HEMBA1000634//Homo sapiens T-cell activation protein (PGR1) gene, complete cds.//0//994bp//99%//AF116272
C-HEMBA1000713//Homo sapiens 10kD protein (BC10) mRNA, complete cds.//0//1254bp//99%//AF053470

EP 1 130 094 A2

NASE I).//1.00E-77//359aa//44%//Q14012
 C-Y79AA1001013
 C-Y79AA1001056//Homo sapiens MAID protein mRNA, complete cds.//0//1475bp//99%//AF113535
 C-Y79AA1001062//TUMOR NECROSIS FACTOR, ALPHA-INDUCED PROTEIN 1, ENDOTHELIAL (B12 PRO-
 5 TEIN).//8.90E-12//132aa//38%//Q13829
 C-Y79AA1001090//NUCLEAR FACTOR NF-KAPPA-B P105 SUBUNIT (DNA-BINDING FACTOR KBF1) (EBP- 1)
 (NF-KAPPA-B1 P84 /NF-KAPPA-B1 P98) [CONTAINS: NUCLEAR FACTOR NF- KAPPA-B P50 SUBUNIT] (FRAG-
 MENT).//4.50E-09//144aa//31%//Q63369
 C-Y79AA1001212//Homo sapiens SL15 protein mRNA, complete cds.//6.30E-306//1388bp//99%//AF038961
 10 C-Y79AA1001264//HYPOTHETICAL 39.9 KD PROTEIN T15H9.1 IN CHROMOSOME II PRECURSOR.//5.10E-
 106//351aa//58%//Q10005
 C-Y79AA1001272//Homo sapiens retinoic acid repressible protein (RARG-1) mRNA, complete cds.//1.50E-183//
 867bp//98%//AF172066
 C-Y79AA1001328//Mus musculus mRNA for Dll3 protein, complete cds.//1.90E-263//1988bp//79%//AB013440
 15 C-Y79AA1001426//ANION EXCHANGE PROTEIN 3 (CARDIAC/BRAIN BAND 3-LIKE PROTEIN) (CAE3/BAE3).//
 6.20E-66//609aa//31%//P48751
 C-Y79AA1001427//Homo sapiens cytochrome b5 reductase 1 (B5R.1) mRNA, complete cds.//0//1588bp//99%//
 AF169481
 C-Y79AA1001430//Homo sapiens mRNA for KIAA0469 protein, complete cds.//0//2943bp//99%//AB007938
 20 C-Y79AA1001523//Homo sapiens transcriptional intermediary factor 1 alpha mRNA, complete cds.//0//2263bp//
 99%//AF119042
 C-Y79AA1001530//Human beta-tubulin gene (5-beta) with ten Alu family members.//0//1920bp//98%//X00734
 C-Y79AA1001592
 C-Y79AA1001727//CELL SURFACE A33 ANTIGEN PRECURSOR.//1.10E-13//286aa//27%//Q99795
 25 C-Y79AA1001787//PROBABLE CALCIUM-TRANSPORTING ATPASE 9 (EC 3.6. 1.38).//1.70E-133//544aa//37%//
 Q12697
 C-Y79AA1001793//Mus musculus mRNA for GSG1, complete cds.//3.70E-126//532bp//78%//D87325
 C-Y79AA1001795//Homo sapiens mRNA for GalT4 protein.//2.30E-250//1137bp//99%//Y15061
 C-Y79AA1001799//MITOCHONDRIAL RNA SPLICING PROTEIN MSR4.//3.40E-54//182aa//39%//P23500
 30 C-Y79AA1001803//Homo sapiens secretogranin III mRNA, complete cds.//0//1871bp//99%//AF078851
 C-Y79AA1001863
 C-Y79AA1002022//POLIOVIRUS RECEPTOR HOMOLOG PRECURSOR.//2.20E-06//140aa//26%//P32507
 C-Y79AA1002058//Mus musculus Gng3lg mRNA, complete cds.//4.10E-167//1145bp//83%//AF069954
 C-Y79AA1002121//HISTONE H1.//4.90E-12//114aa//35%//P35060
 35 C-Y79AA1002129
 C-Y79AA1002213//HYPOTHETICAL 52.7 KD PROTEIN C38C10.2 IN CHROMOSOME III.//1.20E-98//262aa//
 41%//Q03567
 C-Y79AA1002334//GLUCOSE REPRESSION MEDIATOR PROTEIN.//1.70E-10//333aa//23%//P14922
 C-Y79AA1002373//Mus musculus mRNA for GSG1, complete cds.//7.20E-147//680bp//79%//D87325
 40 C-Y79AA1002376//Rattus norvegicus cytoplasmic dynein intermediate chain 2B mRNA, complete cds.//1.50E-
 304//1667bp//90%//U39045
 C-Y79AA1002378//Homo sapiens zinc finger protein NY-REN-21 antigen mRNA, partial cds.//0//963bp//99%//
 AF155100
 C-Y79AA1002381//Homo sapiens cell cycle related kinase mRNA, complete cds.//0//1791bp//98%//AF035013
 45

Claims

1. Use of an oligonucleotide as a primer for synthesizing the polynucleotide comprising the nucleotide sequence set
 50 forth in any one of SEQ ID NOs: 1-829 and 2545, or the complementary strand thereof, wherein said oligonucleotide
 is complementary to said polynucleotide or the complementary strand thereof and comprises at least 15 nucle-
 otides.
2. A primer set for synthesizing polynucleotides, the primer set comprising an oligo-dT primer and an oligonucleotide
 55 complementary to the complementary strand of the polynucleotide comprising the nucleotide sequence set forth
 in any one of SEQ ID NOs: 1-829 and 2545, wherein said oligonucleotide comprises at least 15 nucleotides.
3. A primer set for synthesizing polynucleotides, the primer set comprising a combination of an oligonucleotide com-

EP 1 130 094 A2

prising a nucleotide sequence complementary to the complementary strand of the polynucleotide comprising a 5'-end nucleotide sequence and an oligonucleotide comprising a nucleotide sequence complementary to the polynucleotide comprising a 3'-end nucleotide sequence, wherein said oligonucleotides comprise at least 15 nucleotides and wherein said combination of 5'-end nucleotide sequence.3'-end nucleotide sequence is selected from the group consisting of:

5

10

15

20

25

30

35

40

45

50

55

SEQ ID NO:4 and SEQ ID NO:830
SEQ ID NO:5 and SEQ ID NO:831
SEQ ID NO:6 and SEQ ID NO:832
SEQ ID NO:7 and SEQ ID NO:833
SEQ ID NO:8 and SEQ ID NO:834
SEQ ID NO:9 and SEQ ID NO:835
SEQ ID NO:11 and SEQ ID NO:836
SEQ ID NO:12 and SEQ ID NO:837
SEQ ID NO:13 and SEQ ID NO:838
SEQ ID NO:14 and SEQ ID NO:839
SEQ ID NO:15 and SEQ ID NO:840
SEQ ID NO:16 and SEQ ID NO:841
SEQ ID NO:17 and SEQ ID NO:842
SEQ ID NO:18 and SEQ ID NO:843
SEQ ID NO:20 and SEQ ID NO:844
SEQ ID NO:22 and SEQ ID NO:845
SEQ ID NO:23 and SEQ ID NO:846
SEQ ID NO:24 and SEQ ID NO:847
SEQ ID NO:25 and SEQ ID NO:848
SEQ ID NO:26 and SEQ ID NO:849
SEQ ID NO:27 and SEQ ID NO:850
SEQ ID NO:28 and SEQ ID NO:851
SEQ ID NO:29 and SEQ ID NO:852
SEQ ID NO:30 and SEQ ID NO:853
SEQ ID NO:31 and SEQ ID NO:854
SEQ ID NO:32 and SEQ ID NO:855
SEQ ID NO:33 and SEQ ID NO:856
SEQ ID NO:34 and SEQ ID NO:857
SEQ ID NO:35 and SEQ ID NO:858
SEQ ID NO:36 and SEQ ID NO:859
SEQ ID NO:37 and SEQ ID NO:860
SEQ ID NO:39 and SEQ ID NO:861
SEQ ID NO:40 and SEQ ID NO:862
SEQ ID NO:41 and SEQ ID NO:863
SEQ ID NO:42 and SEQ ID NO:864
SEQ ID NO:44 and SEQ ID NO:865
SEQ ID NO:45 and SEQ ID NO:866
SEQ ID NO:46 and SEQ ID NO:867
SEQ ID NO:47 and SEQ ID NO:868

EP 1 130 094 A2

5 SEQ ID NO:48 and SEQ ID NO:869
SEQ ID NO:49 and SEQ ID NO:870
SEQ ID NO:50 and SEQ ID NO:871
SEQ ID NO:52 and SEQ ID NO:872
SEQ ID NO:53 and SEQ ID NO:873
SEQ ID NO:54 and SEQ ID NO:874
SEQ ID NO:55 and SEQ ID NO:875
10 SEQ ID NO:56 and SEQ ID NO:876
SEQ ID NO:57 and SEQ ID NO:877
SEQ ID NO:58 and SEQ ID NO:878
SEQ ID NO:59 and SEQ ID NO:879
SEQ ID NO:60 and SEQ ID NO:880
15 SEQ ID NO:61 and SEQ ID NO:881
SEQ ID NO:62 and SEQ ID NO:882
SEQ ID NO:63 and SEQ ID NO:883
SEQ ID NO:64 and SEQ ID NO:884
SEQ ID NO:65 and SEQ ID NO:885
20 SEQ ID NO:66 and SEQ ID NO:886
SEQ ID NO:67 and SEQ ID NO:887
SEQ ID NO:69 and SEQ ID NO:888
SEQ ID NO:70 and SEQ ID NO:889
SEQ ID NO:71 and SEQ ID NO:890
25 SEQ ID NO:72 and SEQ ID NO:891
SEQ ID NO:73 and SEQ ID NO:892
SEQ ID NO:74 and SEQ ID NO:893
SEQ ID NO:75 and SEQ ID NO:894
SEQ ID NO:76 and SEQ ID NO:895
30 SEQ ID NO:77 and SEQ ID NO:896
SEQ ID NO:78 and SEQ ID NO:897
SEQ ID NO:79 and SEQ ID NO:898
SEQ ID NO:80 and SEQ ID NO:899
SEQ ID NO:81 and SEQ ID NO:900
35 SEQ ID NO:82 and SEQ ID NO:901
SEQ ID NO:83 and SEQ ID NO:902
SEQ ID NO:84 and SEQ ID NO:903
SEQ ID NO:85 and SEQ ID NO:904
40 SEQ ID NO:86 and SEQ ID NO:905
SEQ ID NO:87 and SEQ ID NO:906
SEQ ID NO:88 and SEQ ID NO:907
SEQ ID NO:89 and SEQ ID NO:908
SEQ ID NO:90 and SEQ ID NO:909
45 SEQ ID NO:91 and SEQ ID NO:910
SEQ ID NO:92 and SEQ ID NO:911
SEQ ID NO:93 and SEQ ID NO:912
SEQ ID NO:94 and SEQ ID NO:913
SEQ ID NO:95 and SEQ ID NO:914
50 SEQ ID NO:96 and SEQ ID NO:915
SEQ ID NO:97 and SEQ ID NO:916
SEQ ID NO:98 and SEQ ID NO:917
SEQ ID NO:99 and SEQ ID NO:918
55 SEQ ID NO:100 and SEQ ID NO:919
SEQ ID NO:101 and SEQ ID NO:920

EP 1 130 094 A2

5 SEQ ID NO:102 and SEQ ID NO:921
SEQ ID NO:103 and SEQ ID NO:922
SEQ ID NO:104 and SEQ ID NO:923
SEQ ID NO:105 and SEQ ID NO:924
SEQ ID NO:106 and SEQ ID NO:925
SEQ ID NO:107 and SEQ ID NO:926
SEQ ID NO:108 and SEQ ID NO:927
10 SEQ ID NO:109 and SEQ ID NO:928
SEQ ID NO:110 and SEQ ID NO:929
SEQ ID NO:111 and SEQ ID NO:930
SEQ ID NO:112 and SEQ ID NO:931
SEQ ID NO:113 and SEQ ID NO:932
15 SEQ ID NO:114 and SEQ ID NO:933
SEQ ID NO:115 and SEQ ID NO:934
SEQ ID NO:116 and SEQ ID NO:935
SEQ ID NO:117 and SEQ ID NO:936
SEQ ID NO:118 and SEQ ID NO:937
20 SEQ ID NO:119 and SEQ ID NO:938
SEQ ID NO:120 and SEQ ID NO:939
SEQ ID NO:122 and SEQ ID NO:940
SEQ ID NO:123 and SEQ ID NO:941
SEQ ID NO:124 and SEQ ID NO:942
25 SEQ ID NO:125 and SEQ ID NO:943
SEQ ID NO:126 and SEQ ID NO:944
SEQ ID NO:127 and SEQ ID NO:945
SEQ ID NO:128 and SEQ ID NO:946
SEQ ID NO:129 and SEQ ID NO:947
30 SEQ ID NO:130 and SEQ ID NO:948
SEQ ID NO:131 and SEQ ID NO:949
SEQ ID NO:132 and SEQ ID NO:950
SEQ ID NO:133 and SEQ ID NO:951
SEQ ID NO:134 and SEQ ID NO:952
35 SEQ ID NO:135 and SEQ ID NO:953
SEQ ID NO:136 and SEQ ID NO:954
SEQ ID NO:137 and SEQ ID NO:955
SEQ ID NO:138 and SEQ ID NO:956
SEQ ID NO:139 and SEQ ID NO:957
40 SEQ ID NO:140 and SEQ ID NO:958
SEQ ID NO:141 and SEQ ID NO:959
SEQ ID NO:142 and SEQ ID NO:960
SEQ ID NO:143 and SEQ ID NO:961
45 SEQ ID NO:144 and SEQ ID NO:962
SEQ ID NO:145 and SEQ ID NO:963
SEQ ID NO:146 and SEQ ID NO:964
SEQ ID NO:147 and SEQ ID NO:965
SEQ ID NO:148 and SEQ ID NO:966
50 SEQ ID NO:149 and SEQ ID NO:967
SEQ ID NO:150 and SEQ ID NO:968
SEQ ID NO:151 and SEQ ID NO:969
SEQ ID NO:152 and SEQ ID NO:970
55 SEQ ID NO:153 and SEQ ID NO:971
SEQ ID NO:154 and SEQ ID NO:972

EP 1 130 094 A2

5 SEQ ID NO:155 and SEQ ID NO:973
SEQ ID NO:156 and SEQ ID NO:974
SEQ ID NO:157 and SEQ ID NO:975
SEQ ID NO:158 and SEQ ID NO:976
SEQ ID NO:159 and SEQ ID NO:977
SEQ ID NO:160 and SEQ ID NO:978
SEQ ID NO:161 and SEQ ID NO:979
10 SEQ ID NO:162 and SEQ ID NO:980
SEQ ID NO:163 and SEQ ID NO:981
SEQ ID NO:164 and SEQ ID NO:982
SEQ ID NO:165 and SEQ ID NO:983
SEQ ID NO:166 and SEQ ID NO:984
15 SEQ ID NO:167 and SEQ ID NO:985
SEQ ID NO:168 and SEQ ID NO:986
SEQ ID NO:169 and SEQ ID NO:987
SEQ ID NO:170 and SEQ ID NO:988
SEQ ID NO:171 and SEQ ID NO:989
20 SEQ ID NO:172 and SEQ ID NO:990
SEQ ID NO:173 and SEQ ID NO:991
SEQ ID NO:174 and SEQ ID NO:992
SEQ ID NO:175 and SEQ ID NO:993
SEQ ID NO:176 and SEQ ID NO:994
25 SEQ ID NO:177 and SEQ ID NO:995
SEQ ID NO:178 and SEQ ID NO:996
SEQ ID NO:179 and SEQ ID NO:997
SEQ ID NO:180 and SEQ ID NO:998
SEQ ID NO:181 and SEQ ID NO:999
30 SEQ ID NO:182 and SEQ ID NO:1000
SEQ ID NO:183 and SEQ ID NO:1001
SEQ ID NO:184 and SEQ ID NO:1002
SEQ ID NO:185 and SEQ ID NO:1003
35 SEQ ID NO:186 and SEQ ID NO:1004
SEQ ID NO:187 and SEQ ID NO:1005
SEQ ID NO:188 and SEQ ID NO:1006
SEQ ID NO:189 and SEQ ID NO:1007
SEQ ID NO:190 and SEQ ID NO:1008
40 SEQ ID NO:191 and SEQ ID NO:1009
SEQ ID NO:193 and SEQ ID NO:1010
SEQ ID NO:194 and SEQ ID NO:1011
SEQ ID NO:195 and SEQ ID NO:1012
SEQ ID NO:196 and SEQ ID NO:1013
45 SEQ ID NO:197 and SEQ ID NO:1014
SEQ ID NO:198 and SEQ ID NO:1015
SEQ ID NO:199 and SEQ ID NO:1016
SEQ ID NO:200 and SEQ ID NO:1017
SEQ ID NO:201 and SEQ ID NO:1018
50 SEQ ID NO:202 and SEQ ID NO:1019
SEQ ID NO:203 and SEQ ID NO:1020
SEQ ID NO:204 and SEQ ID NO:1021
SEQ ID NO:205 and SEQ ID NO:1022
55 SEQ ID NO:207 and SEQ ID NO:1023
SEQ ID NO:208 and SEQ ID NO:1024

EP 1 130 094 A2

5 SEQ ID NO:209 and SEQ ID NO:1025
SEQ ID NO:210 and SEQ ID NO:1026
SEQ ID NO:211 and SEQ ID NO:1027
SEQ ID NO:212 and SEQ ID NO:1028
SEQ ID NO:213 and SEQ ID NO:1029
SEQ ID NO:214 and SEQ ID NO:1030
SEQ ID NO:215 and SEQ ID NO:1031
10 SEQ ID NO:216 and SEQ ID NO:1032
SEQ ID NO:248 and SEQ ID NO:1033
SEQ ID NO:249 and SEQ ID NO:1034
SEQ ID NO:250 and SEQ ID NO:1035
SEQ ID NO:251 and SEQ ID NO:1036
15 SEQ ID NO:252 and SEQ ID NO:1037
SEQ ID NO:253 and SEQ ID NO:1038
SEQ ID NO:254 and SEQ ID NO:1039
SEQ ID NO:255 and SEQ ID NO:1040
SEQ ID NO:256 and SEQ ID NO:1041
20 SEQ ID NO:257 and SEQ ID NO:1042
SEQ ID NO:258 and SEQ ID NO:1043
SEQ ID NO:259 and SEQ ID NO:1044
SEQ ID NO:260 and SEQ ID NO:1045
SEQ ID NO:261 and SEQ ID NO:1046
25 SEQ ID NO:262 and SEQ ID NO:1047
SEQ ID NO:263 and SEQ ID NO:1048
SEQ ID NO:264 and SEQ ID NO:1049
SEQ ID NO:265 and SEQ ID NO:1050
SEQ ID NO:266 and SEQ ID NO:1051
30 SEQ ID NO:267 and SEQ ID NO:1052
SEQ ID NO:268 and SEQ ID NO:1053
SEQ ID NO:291 and SEQ ID NO:1054
SEQ ID NO:292 and SEQ ID NO:1055
35 SEQ ID NO:293 and SEQ ID NO:1056
SEQ ID NO:294 and SEQ ID NO:1057
SEQ ID NO:295 and SEQ ID NO:1058
SEQ ID NO:296 and SEQ ID NO:1059
SEQ ID NO:297 and SEQ ID NO:1060
40 SEQ ID NO:298 and SEQ ID NO:1061
SEQ ID NO:300 and SEQ ID NO:1062
SEQ ID NO:301 and SEQ ID NO:1063
SEQ ID NO:302 and SEQ ID NO:1064
SEQ ID NO:304 and SEQ ID NO:1065
45 SEQ ID NO:305 and SEQ ID NO:1066
SEQ ID NO:306 and SEQ ID NO:1067
SEQ ID NO:307 and SEQ ID NO:1068
SEQ ID NO:308 and SEQ ID NO:1069
SEQ ID NO:309 and SEQ ID NO:1070
50 SEQ ID NO:310 and SEQ ID NO:1071
SEQ ID NO:311 and SEQ ID NO:1072
SEQ ID NO:312 and SEQ ID NO:1073
SEQ ID NO:313 and SEQ ID NO:1074
55 SEQ ID NO:314 and SEQ ID NO:1075
SEQ ID NO:315 and SEQ ID NO:1076

EP 1 130 094 A2

5 SEQ ID NO:317 and SEQ ID NO:1077
SEQ ID NO:318 and SEQ ID NO:1078
SEQ ID NO:319 and SEQ ID NO:1079
SEQ ID NO:320 and SEQ ID NO:1080
SEQ ID NO:321 and SEQ ID NO:1081
SEQ ID NO:322 and SEQ ID NO:1082
SEQ ID NO:323 and SEQ ID NO:1083
10 SEQ ID NO:324 and SEQ ID NO:1084
SEQ ID NO:325 and SEQ ID NO:1085
SEQ ID NO:326 and SEQ ID NO:1086
SEQ ID NO:327 and SEQ ID NO:1087
SEQ ID NO:328 and SEQ ID NO:1088
15 SEQ ID NO:329 and SEQ ID NO:1089
SEQ ID NO:330 and SEQ ID NO:1090
SEQ ID NO:331 and SEQ ID NO:1091
SEQ ID NO:332 and SEQ ID NO:1092
SEQ ID NO:333 and SEQ ID NO:1093
20 SEQ ID NO:334 and SEQ ID NO:1094
SEQ ID NO:335 and SEQ ID NO:1095
SEQ ID NO:336 and SEQ ID NO:1096
SEQ ID NO:337 and SEQ ID NO:1097
SEQ ID NO:338 and SEQ ID NO:1098
25 SEQ ID NO:339 and SEQ ID NO:1099
SEQ ID NO:340 and SEQ ID NO:1100
SEQ ID NO:341 and SEQ ID NO:1101
SEQ ID NO:342 and SEQ ID NO:1102
SEQ ID NO:343 and SEQ ID NO:1103
30 SEQ ID NO:344 and SEQ ID NO:1104
SEQ ID NO:345 and SEQ ID NO:1105
SEQ ID NO:347 and SEQ ID NO:1106
SEQ ID NO:348 and SEQ ID NO:1107
SEQ ID NO:349 and SEQ ID NO:1108
35 SEQ ID NO:350 and SEQ ID NO:1109
SEQ ID NO:351 and SEQ ID NO:1110
SEQ ID NO:352 and SEQ ID NO:1111
SEQ ID NO:353 and SEQ ID NO:1112
40 SEQ ID NO:354 and SEQ ID NO:1113
SEQ ID NO:355 and SEQ ID NO:1114
SEQ ID NO:356 and SEQ ID NO:1115
SEQ ID NO:357 and SEQ ID NO:1116
SEQ ID NO:358 and SEQ ID NO:1117
45 SEQ ID NO:359 and SEQ ID NO:1118
SEQ ID NO:360 and SEQ ID NO:1119
SEQ ID NO:361 and SEQ ID NO:1120
SEQ ID NO:362 and SEQ ID NO:1121
SEQ ID NO:363 and SEQ ID NO:1122
50 SEQ ID NO:364 and SEQ ID NO:1123
SEQ ID NO:365 and SEQ ID NO:1124
SEQ ID NO:366 and SEQ ID NO:1125
SEQ ID NO:367 and SEQ ID NO:1126
55 SEQ ID NO:368 and SEQ ID NO:1127
SEQ ID NO:369 and SEQ ID NO:1128

EP 1 130 094 A2

5 SEQ ID NO:370 and SEQ ID NO:1129
SEQ ID NO:371 and SEQ ID NO:1130
SEQ ID NO:372 and SEQ ID NO:1131
SEQ ID NO:373 and SEQ ID NO:1132
SEQ ID NO:374 and SEQ ID NO:1133
SEQ ID NO:375 and SEQ ID NO:1134
SEQ ID NO:376 and SEQ ID NO:1135
10 SEQ ID NO:377 and SEQ ID NO:1136
SEQ ID NO:378 and SEQ ID NO:1137
SEQ ID NO:379 and SEQ ID NO:1138
SEQ ID NO:380 and SEQ ID NO:1139
SEQ ID NO:381 and SEQ ID NO:1140
15 SEQ ID NO:382 and SEQ ID NO:1141
SEQ ID NO:383 and SEQ ID NO:1142
SEQ ID NO:384 and SEQ ID NO:1143
SEQ ID NO:385 and SEQ ID NO:1144
SEQ ID NO:386 and SEQ ID NO:1145
20 SEQ ID NO:387 and SEQ ID NO:1146
SEQ ID NO:388 and SEQ ID NO:1147
SEQ ID NO:389 and SEQ ID NO:1148
SEQ ID NO:390 and SEQ ID NO:1149
SEQ ID NO:391 and SEQ ID NO:1150
25 SEQ ID NO:392 and SEQ ID NO:1151
SEQ ID NO:393 and SEQ ID NO:1152
SEQ ID NO:394 and SEQ ID NO:1153
SEQ ID NO:395 and SEQ ID NO:1154
SEQ ID NO:396 and SEQ ID NO:1155
30 SEQ ID NO:397 and SEQ ID NO:1156
SEQ ID NO:398 and SEQ ID NO:1157
SEQ ID NO:399 and SEQ ID NO:1158
SEQ ID NO:400 and SEQ ID NO:1159
35 SEQ ID NO:401 and SEQ ID NO:1160
SEQ ID NO:402 and SEQ ID NO:1161
SEQ ID NO:403 and SEQ ID NO:1162
SEQ ID NO:404 and SEQ ID NO:1163
SEQ ID NO:405 and SEQ ID NO:1164
40 SEQ ID NO:406 and SEQ ID NO:1165
SEQ ID NO:407 and SEQ ID NO:1166
SEQ ID NO:408 and SEQ ID NO:1167
SEQ ID NO:409 and SEQ ID NO:1168
SEQ ID NO:410 and SEQ ID NO:1169
45 SEQ ID NO:411 and SEQ ID NO:1170
SEQ ID NO:412 and SEQ ID NO:1171
SEQ ID NO:413 and SEQ ID NO:1172
SEQ ID NO:414 and SEQ ID NO:1173
SEQ ID NO:415 and SEQ ID NO:1174
50 SEQ ID NO:416 and SEQ ID NO:1175
SEQ ID NO:417 and SEQ ID NO:1176
SEQ ID NO:419 and SEQ ID NO:1177
SEQ ID NO:420 and SEQ ID NO:1178
55 SEQ ID NO:421 and SEQ ID NO:1179
SEQ ID NO:422 and SEQ ID NO:1180

EP 1 130 094 A2

5 SEQ ID NO:423 and SEQ ID NO:1181
SEQ ID NO:424 and SEQ ID NO:1182
SEQ ID NO:425 and SEQ ID NO:1183
SEQ ID NO:426 and SEQ ID NO:1184
SEQ ID NO:427 and SEQ ID NO:1185
SEQ ID NO:428 and SEQ ID NO:1186
SEQ ID NO:429 and SEQ ID NO:1187
10 SEQ ID NO:430 and SEQ ID NO:1188
SEQ ID NO:431 and SEQ ID NO:1189
SEQ ID NO:432 and SEQ ID NO:1190
SEQ ID NO:433 and SEQ ID NO:1191
SEQ ID NO:434 and SEQ ID NO:1192
15 SEQ ID NO:435 and SEQ ID NO:1193
SEQ ID NO:436 and SEQ ID NO:1194
SEQ ID NO:437 and SEQ ID NO:1195
SEQ ID NO:438 and SEQ ID NO:1196
SEQ ID NO:439 and SEQ ID NO:1197
20 SEQ ID NO:440 and SEQ ID NO:1198
SEQ ID NO:441 and SEQ ID NO:1199
SEQ ID NO:442 and SEQ ID NO:1200
SEQ ID NO:443 and SEQ ID NO:1201
SEQ ID NO:444 and SEQ ID NO:1202
25 SEQ ID NO:446 and SEQ ID NO:1203
SEQ ID NO:447 and SEQ ID NO:1204
SEQ ID NO:448 and SEQ ID NO:1205
SEQ ID NO:450 and SEQ ID NO:1206
SEQ ID NO:451 and SEQ ID NO:1207
30 SEQ ID NO:452 and SEQ ID NO:1208
SEQ ID NO:453 and SEQ ID NO:1209
SEQ ID NO:454 and SEQ ID NO:1210
SEQ ID NO:455 and SEQ ID NO:1211
35 SEQ ID NO:456 and SEQ ID NO:1212
SEQ ID NO:457 and SEQ ID NO:1213
SEQ ID NO:458 and SEQ ID NO:1214
SEQ ID NO:459 and SEQ ID NO:1215
SEQ ID NO:460 and SEQ ID NO:1216
40 SEQ ID NO:461 and SEQ ID NO:1217
SEQ ID NO:462 and SEQ ID NO:1218
SEQ ID NO:463 and SEQ ID NO:1219
SEQ ID NO:464 and SEQ ID NO:1220
SEQ ID NO:465 and SEQ ID NO:1221
45 SEQ ID NO:466 and SEQ ID NO:1222
SEQ ID NO:467 and SEQ ID NO:1223
SEQ ID NO:468 and SEQ ID NO:1224
SEQ ID NO:469 and SEQ ID NO:1225
SEQ ID NO:470 and SEQ ID NO:1226
50 SEQ ID NO:471 and SEQ ID NO:1227
SEQ ID NO:472 and SEQ ID NO:1228
SEQ ID NO:473 and SEQ ID NO:1229
SEQ ID NO:474 and SEQ ID NO:1230
55 SEQ ID NO:475 and SEQ ID NO:1231
SEQ ID NO:476 and SEQ ID NO:1232

EP 1 130 094 A2

5 SEQ ID NO:477 and SEQ ID NO:1233
SEQ ID NO:478 and SEQ ID NO:1234
SEQ ID NO:479 and SEQ ID NO:1235
SEQ ID NO:480 and SEQ ID NO:1236
SEQ ID NO:481 and SEQ ID NO:1237
SEQ ID NO:482 and SEQ ID NO:1238
SEQ ID NO:483 and SEQ ID NO:1239
10 SEQ ID NO:484 and SEQ ID NO:1240
SEQ ID NO:485 and SEQ ID NO:1241
SEQ ID NO:486 and SEQ ID NO:1242
SEQ ID NO:487 and SEQ ID NO:1243
SEQ ID NO:488 and SEQ ID NO:1244
15 SEQ ID NO:489 and SEQ ID NO:1245
SEQ ID NO:490 and SEQ ID NO:1246
SEQ ID NO:492 and SEQ ID NO:1247
SEQ ID NO:493 and SEQ ID NO:1248
SEQ ID NO:494 and SEQ ID NO:1249
20 SEQ ID NO:496 and SEQ ID NO:1250
SEQ ID NO:497 and SEQ ID NO:1251
SEQ ID NO:498 and SEQ ID NO:1252
SEQ ID NO:499 and SEQ ID NO:1253
SEQ ID NO:500 and SEQ ID NO:1254
25 SEQ ID NO:501 and SEQ ID NO:1255
SEQ ID NO:502 and SEQ ID NO:1256
SEQ ID NO:503 and SEQ ID NO:1257
SEQ ID NO:504 and SEQ ID NO:1258
SEQ ID NO:505 and SEQ ID NO:1259
30 SEQ ID NO:506 and SEQ ID NO:1260
SEQ ID NO:507 and SEQ ID NO:1261
SEQ ID NO:508 and SEQ ID NO:1262
SEQ ID NO:509 and SEQ ID NO:1263
35 SEQ ID NO:510 and SEQ ID NO:1264
SEQ ID NO:511 and SEQ ID NO:1265
SEQ ID NO:513 and SEQ ID NO:1266
SEQ ID NO:514 and SEQ ID NO:1267
SEQ ID NO:515 and SEQ ID NO:1268
40 SEQ ID NO:516 and SEQ ID NO:1269
SEQ ID NO:517 and SEQ ID NO:1270
SEQ ID NO:518 and SEQ ID NO:1271
SEQ ID NO:519 and SEQ ID NO:1272
SEQ ID NO:520 and SEQ ID NO:1273
45 SEQ ID NO:521 and SEQ ID NO:1274
SEQ ID NO:522 and SEQ ID NO:1275
SEQ ID NO:523 and SEQ ID NO:1276
SEQ ID NO:524 and SEQ ID NO:1277
SEQ ID NO:525 and SEQ ID NO:1278
50 SEQ ID NO:526 and SEQ ID NO:1279
SEQ ID NO:527 and SEQ ID NO:1280
SEQ ID NO:528 and SEQ ID NO:1281
SEQ ID NO:529 and SEQ ID NO:1282
55 SEQ ID NO:530 and SEQ ID NO:1283
SEQ ID NO:531 and SEQ ID NO:1284

EP 1 130 094 A2

5 SEQ ID NO:532 and SEQ ID NO:1285
SEQ ID NO:533 and SEQ ID NO:1286
SEQ ID NO:534 and SEQ ID NO:1287
SEQ ID NO:535 and SEQ ID NO:1288
SEQ ID NO:536 and SEQ ID NO:1289
SEQ ID NO:537 and SEQ ID NO:1290
SEQ ID NO:538 and SEQ ID NO:1291
10 SEQ ID NO:539 and SEQ ID NO:1292
SEQ ID NO:540 and SEQ ID NO:1293
SEQ ID NO:541 and SEQ ID NO:1294
SEQ ID NO:542 and SEQ ID NO:1295
SEQ ID NO:543 and SEQ ID NO:1296
15 SEQ ID NO:544 and SEQ ID NO:1297
SEQ ID NO:545 and SEQ ID NO:1298
SEQ ID NO:546 and SEQ ID NO:1299
SEQ ID NO:547 and SEQ ID NO:1300
SEQ ID NO:548 and SEQ ID NO:1301
20 SEQ ID NO:549 and SEQ ID NO:1302
SEQ ID NO:550 and SEQ ID NO:1303
SEQ ID NO:551 and SEQ ID NO:1304
SEQ ID NO:552 and SEQ ID NO:1305
SEQ ID NO:553 and SEQ ID NO:1306
25 SEQ ID NO:554 and SEQ ID NO:1307
SEQ ID NO:555 and SEQ ID NO:1308
SEQ ID NO:556 and SEQ ID NO:1309
SEQ ID NO:557 and SEQ ID NO:1310
SEQ ID NO:558 and SEQ ID NO:1311
30 SEQ ID NO:559 and SEQ ID NO:1312
SEQ ID NO:560 and SEQ ID NO:1313
SEQ ID NO:561 and SEQ ID NO:1314
SEQ ID NO:562 and SEQ ID NO:1315
35 SEQ ID NO:563 and SEQ ID NO:1316
SEQ ID NO:564 and SEQ ID NO:1317
SEQ ID NO:565 and SEQ ID NO:1318
SEQ ID NO:566 and SEQ ID NO:1319
SEQ ID NO:567 and SEQ ID NO:1320
40 SEQ ID NO:569 and SEQ ID NO:1321
SEQ ID NO:570 and SEQ ID NO:1322
SEQ ID NO:571 and SEQ ID NO:1323
SEQ ID NO:572 and SEQ ID NO:1324
SEQ ID NO:573 and SEQ ID NO:1325
45 SEQ ID NO:574 and SEQ ID NO:1326
SEQ ID NO:575 and SEQ ID NO:1327
SEQ ID NO:576 and SEQ ID NO:1328
SEQ ID NO:577 and SEQ ID NO:1329
SEQ ID NO:579 and SEQ ID NO:1330
50 SEQ ID NO:580 and SEQ ID NO:1331
SEQ ID NO:581 and SEQ ID NO:1332
SEQ ID NO:582 and SEQ ID NO:1333
SEQ ID NO:583 and SEQ ID NO:1334
55 SEQ ID NO:584 and SEQ ID NO:1335
SEQ ID NO:585 and SEQ ID NO:1336

EP 1 130 094 A2

5 SEQ ID NO:586 and SEQ ID NO:1337
SEQ ID NO:587 and SEQ ID NO:1338
SEQ ID NO:588 and SEQ ID NO:1339
SEQ ID NO:589 and SEQ ID NO:1340
SEQ ID NO:590 and SEQ ID NO:1341
SEQ ID NO:591 and SEQ ID NO:1342
SEQ ID NO:592 and SEQ ID NO:1343
10 SEQ ID NO:593 and SEQ ID NO:1344
SEQ ID NO:594 and SEQ ID NO:1345
SEQ ID NO:595 and SEQ ID NO:1346
SEQ ID NO:596 and SEQ ID NO:1347
SEQ ID NO:597 and SEQ ID NO:1348
15 SEQ ID NO:598 and SEQ ID NO:1349
SEQ ID NO:599 and SEQ ID NO:1350
SEQ ID NO:600 and SEQ ID NO:1351
SEQ ID NO:601 and SEQ ID NO:1352
SEQ ID NO:602 and SEQ ID NO:1353
20 SEQ ID NO:603 and SEQ ID NO:1354
SEQ ID NO:604 and SEQ ID NO:1355
SEQ ID NO:605 and SEQ ID NO:1356
SEQ ID NO:606 and SEQ ID NO:1357
SEQ ID NO:607 and SEQ ID NO:1358
25 SEQ ID NO:608 and SEQ ID NO:1359
SEQ ID NO:609 and SEQ ID NO:1360
SEQ ID NO:610 and SEQ ID NO:1361
SEQ ID NO:611 and SEQ ID NO:1362
SEQ ID NO:612 and SEQ ID NO:1363
30 SEQ ID NO:613 and SEQ ID NO:1364
SEQ ID NO:614 and SEQ ID NO:1365
SEQ ID NO:615 and SEQ ID NO:1366
SEQ ID NO:616 and SEQ ID NO:1367
35 SEQ ID NO:617 and SEQ ID NO:1368
SEQ ID NO:618 and SEQ ID NO:1369
SEQ ID NO:619 and SEQ ID NO:1370
SEQ ID NO:620 and SEQ ID NO:1371
SEQ ID NO:621 and SEQ ID NO:1372
40 SEQ ID NO:622 and SEQ ID NO:1373
SEQ ID NO:623 and SEQ ID NO:1374
SEQ ID NO:624 and SEQ ID NO:1375
SEQ ID NO:625 and SEQ ID NO:1376
SEQ ID NO:626 and SEQ ID NO:1377
45 SEQ ID NO:627 and SEQ ID NO:1378
SEQ ID NO:628 and SEQ ID NO:1379
SEQ ID NO:629 and SEQ ID NO:1380
SEQ ID NO:630 and SEQ ID NO:1381
50 SEQ ID NO:631 and SEQ ID NO:1382
SEQ ID NO:632 and SEQ ID NO:1383
SEQ ID NO:633 and SEQ ID NO:1384
SEQ ID NO:634 and SEQ ID NO:1385
SEQ ID NO:635 and SEQ ID NO:1386
55 SEQ ID NO:636 and SEQ ID NO:1387
SEQ ID NO:637 and SEQ ID NO:1388

EP 1 130 094 A2

5 SEQ ID NO:638 and SEQ ID NO:1389
SEQ ID NO:639 and SEQ ID NO:1390
SEQ ID NO:640 and SEQ ID NO:1391
SEQ ID NO:641 and SEQ ID NO:1392
SEQ ID NO:642 and SEQ ID NO:1393
SEQ ID NO:643 and SEQ ID NO:1394
SEQ ID NO:644 and SEQ ID NO:1395
10 SEQ ID NO:645 and SEQ ID NO:1396
SEQ ID NO:646 and SEQ ID NO:1397
SEQ ID NO:647 and SEQ ID NO:1398
SEQ ID NO:648 and SEQ ID NO:1399
SEQ ID NO:649 and SEQ ID NO:1400
15 SEQ ID NO:650 and SEQ ID NO:1401
SEQ ID NO:651 and SEQ ID NO:1402
SEQ ID NO:652 and SEQ ID NO:1403
SEQ ID NO:653 and SEQ ID NO:1404
SEQ ID NO:654 and SEQ ID NO:1405
20 SEQ ID NO:655 and SEQ ID NO:1406
SEQ ID NO:656 and SEQ ID NO:1407
SEQ ID NO:657 and SEQ ID NO:1408
SEQ ID NO:658 and SEQ ID NO:1409
SEQ ID NO:659 and SEQ ID NO:1410
25 SEQ ID NO:660 and SEQ ID NO:1411
SEQ ID NO:661 and SEQ ID NO:1412
SEQ ID NO:662 and SEQ ID NO:1413
SEQ ID NO:663 and SEQ ID NO:1414
SEQ ID NO:664 and SEQ ID NO:1415
30 SEQ ID NO:665 and SEQ ID NO:1416
SEQ ID NO:666 and SEQ ID NO:1417
SEQ ID NO:667 and SEQ ID NO:1418
SEQ ID NO:668 and SEQ ID NO:1419
35 SEQ ID NO:669 and SEQ ID NO:1420
SEQ ID NO:670 and SEQ ID NO:1421
SEQ ID NO:671 and SEQ ID NO:1422
SEQ ID NO:672 and SEQ ID NO:1423
SEQ ID NO:673 and SEQ ID NO:1424
40 SEQ ID NO:674 and SEQ ID NO:1425
SEQ ID NO:675 and SEQ ID NO:1426
SEQ ID NO:676 and SEQ ID NO:1427
SEQ ID NO:677 and SEQ ID NO:1428
SEQ ID NO:678 and SEQ ID NO:1429
45 SEQ ID NO:679 and SEQ ID NO:1430
SEQ ID NO:680 and SEQ ID NO:1431
SEQ ID NO:681 and SEQ ID NO:1432
SEQ ID NO:682 and SEQ ID NO:1433
50 SEQ ID NO:683 and SEQ ID NO:1434
SEQ ID NO:684 and SEQ ID NO:1435
SEQ ID NO:685 and SEQ ID NO:1436
SEQ ID NO:686 and SEQ ID NO:1437
SEQ ID NO:687 and SEQ ID NO:1438
55 SEQ ID NO:688 and SEQ ID NO:1439
SEQ ID NO:689 and SEQ ID NO:1440

EP 1 130 094 A2

5 SEQ ID NO:690 and SEQ ID NO:1441
SEQ ID NO:691 and SEQ ID NO:1442
SEQ ID NO:692 and SEQ ID NO:1443
SEQ ID NO:693 and SEQ ID NO:1444
SEQ ID NO:694 and SEQ ID NO:1445
SEQ ID NO:695 and SEQ ID NO:1446
SEQ ID NO:696 and SEQ ID NO:1447
10 SEQ ID NO:697 and SEQ ID NO:1448
SEQ ID NO:698 and SEQ ID NO:1449
SEQ ID NO:699 and SEQ ID NO:1450
SEQ ID NO:700 and SEQ ID NO:1451
SEQ ID NO:701 and SEQ ID NO:1452
15 SEQ ID NO:702 and SEQ ID NO:1453
SEQ ID NO:703 and SEQ ID NO:1454
SEQ ID NO:704 and SEQ ID NO:1455
SEQ ID NO:705 and SEQ ID NO:1456
SEQ ID NO:706 and SEQ ID NO:1457
20 SEQ ID NO:707 and SEQ ID NO:1458
SEQ ID NO:708 and SEQ ID NO:1459
SEQ ID NO:709 and SEQ ID NO:1460
SEQ ID NO:710 and SEQ ID NO:1461
SEQ ID NO:711 and SEQ ID NO:1462
25 SEQ ID NO:712 and SEQ ID NO:1463
SEQ ID NO:713 and SEQ ID NO:1464
SEQ ID NO:714 and SEQ ID NO:1465
SEQ ID NO:715 and SEQ ID NO:1466
SEQ ID NO:716 and SEQ ID NO:1467
30 SEQ ID NO:717 and SEQ ID NO:1468
SEQ ID NO:718 and SEQ ID NO:1469
SEQ ID NO:719 and SEQ ID NO:1470
SEQ ID NO:720 and SEQ ID NO:1471
35 SEQ ID NO:721 and SEQ ID NO:1472
SEQ ID NO:722 and SEQ ID NO:1473
SEQ ID NO:723 and SEQ ID NO:1474
SEQ ID NO:724 and SEQ ID NO:1475
SEQ ID NO:725 and SEQ ID NO:1476
40 SEQ ID NO:726 and SEQ ID NO:1477
SEQ ID NO:727 and SEQ ID NO:1478
SEQ ID NO:729 and SEQ ID NO:1479
SEQ ID NO:730 and SEQ ID NO:1480
SEQ ID NO:731 and SEQ ID NO:1481
45 SEQ ID NO:732 and SEQ ID NO:1482
SEQ ID NO:733 and SEQ ID NO:1483
SEQ ID NO:734 and SEQ ID NO:1484
SEQ ID NO:735 and SEQ ID NO:1485
SEQ ID NO:736 and SEQ ID NO:1486
50 SEQ ID NO:737 and SEQ ID NO:1487
SEQ ID NO:738 and SEQ ID NO:1488
SEQ ID NO:739 and SEQ ID NO:1489
SEQ ID NO:740 and SEQ ID NO:1490
55 SEQ ID NO:741 and SEQ ID NO:1491
SEQ ID NO:745 and SEQ ID NO:1492

EP 1 130 094 A2

5 SEQ ID NO:746 and SEQ ID NO:1493
SEQ ID NO:747 and SEQ ID NO:1494
SEQ ID NO:748 and SEQ ID NO:1495
SEQ ID NO:749 and SEQ ID NO:1496
SEQ ID NO:750 and SEQ ID NO:1497
SEQ ID NO:751 and SEQ ID NO:1498
SEQ ID NO:752 and SEQ ID NO:1499
10 SEQ ID NO:753 and SEQ ID NO:1500
SEQ ID NO:754 and SEQ ID NO:1501
SEQ ID NO:755 and SEQ ID NO:1502
SEQ ID NO:756 and SEQ ID NO:1503
SEQ ID NO:757 and SEQ ID NO:1504
15 SEQ ID NO:758 and SEQ ID NO:1505
SEQ ID NO:759 and SEQ ID NO:1506
SEQ ID NO:760 and SEQ ID NO:1507
SEQ ID NO:761 and SEQ ID NO:1508
SEQ ID NO:762 and SEQ ID NO:1509
20 SEQ ID NO:763 and SEQ ID NO:1510
SEQ ID NO:764 and SEQ ID NO:1511
SEQ ID NO:765 and SEQ ID NO:1512
SEQ ID NO:766 and SEQ ID NO:1513
SEQ ID NO:767 and SEQ ID NO:1514
25 SEQ ID NO:768 and SEQ ID NO:1515
SEQ ID NO:769 and SEQ ID NO:1516
SEQ ID NO:770 and SEQ ID NO:1517
SEQ ID NO:771 and SEQ ID NO:1518
SEQ ID NO:772 and SEQ ID NO:1519
30 SEQ ID NO:773 and SEQ ID NO:1520
SEQ ID NO:774 and SEQ ID NO:1521
SEQ ID NO:775 and SEQ ID NO:1522
SEQ ID NO:776 and SEQ ID NO:1523
SEQ ID NO:777 and SEQ ID NO:1524
35 SEQ ID NO:778 and SEQ ID NO:1525
SEQ ID NO:779 and SEQ ID NO:1526
SEQ ID NO:780 and SEQ ID NO:1527
SEQ ID NO:781 and SEQ ID NO:1528
SEQ ID NO:782 and SEQ ID NO:1529
40 SEQ ID NO:783 and SEQ ID NO:1530
SEQ ID NO:784 and SEQ ID NO:1531
SEQ ID NO:785 and SEQ ID NO:1532
SEQ ID NO:786 and SEQ ID NO:1533
45 SEQ ID NO:787 and SEQ ID NO:1534
SEQ ID NO:788 and SEQ ID NO:1535
SEQ ID NO:789 and SEQ ID NO:1536
SEQ ID NO:790 and SEQ ID NO:1537
SEQ ID NO:791 and SEQ ID NO:1538
50 SEQ ID NO:792 and SEQ ID NO:1539
SEQ ID NO:793 and SEQ ID NO:1540
SEQ ID NO:794 and SEQ ID NO:1541
SEQ ID NO:795 and SEQ ID NO:1542
SEQ ID NO:797 and SEQ ID NO:1543
55 SEQ ID NO:798 and SEQ ID NO:1544

5 SEQ ID NO:799 and SEQ ID NO:1545
 SEQ ID NO:800 and SEQ ID NO:1546
 SEQ ID NO:801 and SEQ ID NO:1547
 SEQ ID NO:802 and SEQ ID NO:1548
 SEQ ID NO:803 and SEQ ID NO:1549
 SEQ ID NO:804 and SEQ ID NO:1550
 10 SEQ ID NO:805 and SEQ ID NO:1551
 SEQ ID NO:806 and SEQ ID NO:1552
 SEQ ID NO:807 and SEQ ID NO:1553
 SEQ ID NO:809 and SEQ ID NO:1554
 SEQ ID NO:810 and SEQ ID NO:1555
 15 SEQ ID NO:811 and SEQ ID NO:1556
 SEQ ID NO:812 and SEQ ID NO:1557
 SEQ ID NO:813 and SEQ ID NO:1558
 SEQ ID NO:814 and SEQ ID NO:1559
 20 SEQ ID NO:816 and SEQ ID NO:1560
 SEQ ID NO:817 and SEQ ID NO:1561
 SEQ ID NO:818 and SEQ ID NO:1562
 SEQ ID NO:819 and SEQ ID NO:1563
 25 SEQ ID NO:820 and SEQ ID NO:1564
 SEQ ID NO:822 and SEQ ID NO:1565
 SEQ ID NO:823 and SEQ ID NO:1566
 SEQ ID NO:824 and SEQ ID NO:1567
 30 SEQ ID NO:825 and SEQ ID NO:1568
 SEQ ID NO:826 and SEQ ID NO:1569
 SEQ ID NO:827 and SEQ ID NO:1570
 SEQ ID NO:828 and SEQ ID NO:1571
 SEQ ID NO:829 and SEQ ID NO:1572, and
 35 SEQ ID NO:2545 and SEQ ID NO:2546

4. A polynucleotide which can be synthesized with the primer set of claim 2 or 3.
5. A polynucleotide comprising a coding region in the polynucleotide of claim 4.
- 40 6. A substantially pure protein encoded by polynucleotide of claim 4.
7. A partial peptide of the protein of claim 6.
- 45 8. An isolated polynucleotide selected from the group consisting of
 - (a) a polynucleotide comprising a coding region of the nucleotide sequence set forth in any one of the following SEQ ID NOs:

50

55

EP 1 130 094 A2

5 SEQ ID NO:2547, SEQ ID NO:nnnn , SEQ ID NO:2551, SEQ ID NO:2553, SEQ ID NO:2555,
SEQ ID NO:2557, SEQ ID NO:2558, SEQ ID NO:2560, SEQ ID NO:2562, SEQ ID NO:2564,
SEQ ID NO:2566, SEQ ID NO:2568, SEQ ID NO:2570, SEQ ID NO:2572, SEQ ID NO:2574,
10 SEQ ID NO:2576, SEQ ID NO:2578, SEQ ID NO:2580, SEQ ID NO:2581, SEQ ID NO:2583,
SEQ ID NO:2584, SEQ ID NO:nnnn , SEQ ID NO:2588, SEQ ID NO:2590, SEQ ID NO:2592,
SEQ ID NO:2594, SEQ ID NO:2596, SEQ ID NO:2598, SEQ ID NO:nnnn , SEQ ID NO:2602,
SEQ ID NO:2604, SEQ ID NO:2606, SEQ ID NO:2607, SEQ ID NO:2609, SEQ ID NO:2611,
15 SEQ ID NO:2613, SEQ ID NO:2615, SEQ ID NO:2617, SEQ ID NO:2619, SEQ ID NO:2621,
SEQ ID NO:2623, SEQ ID NO:2625, SEQ ID NO:2627, SEQ ID NO:2629, SEQ ID NO:2630,
SEQ ID NO:2632, SEQ ID NO:2634, SEQ ID NO:2636, SEQ ID NO:2638, SEQ ID NO:2640,
SEQ ID NO:2642, SEQ ID NO:2644, SEQ ID NO:2646, SEQ ID NO:2648, SEQ ID NO:2650,
SEQ ID NO:2652, SEQ ID NO:2654, SEQ ID NO:2656, SEQ ID NO:2658, SEQ ID NO:2660,
20 SEQ ID NO:2662, SEQ ID NO:2664, SEQ ID NO:2666, SEQ ID NO:2668, SEQ ID NO:2670,
SEQ ID NO:2672, SEQ ID NO:2674, SEQ ID NO:2676, SEQ ID NO:2678, SEQ ID NO:2680,
SEQ ID NO:2681, SEQ ID NO:2683, SEQ ID NO:2685, SEQ ID NO:2686, SEQ ID NO:2688,
SEQ ID NO:2690, SEQ ID NO:2692, SEQ ID NO:2694, SEQ ID NO:2696, SEQ ID NO:2698,
SEQ ID NO:2700, SEQ ID NO:2702, SEQ ID NO:2704, SEQ ID NO:2706, SEQ ID NO:2708,
25
30
35
40
45
50
55

EP 1 130 094 A2

5 SEQ ID NO:2710, SEQ ID NO:2712, SEQ ID NO:2714, SEQ ID NO:2715, SEQ ID NO:2717,
 SEQ ID NO:2718, SEQ ID NO:2719, SEQ ID NO:2721, SEQ ID NO:2722, SEQ ID NO:2724,
 SEQ ID NO:2726, SEQ ID NO:2728, SEQ ID NO:2730, SEQ ID NO:2732, SEQ ID NO:2734
 10 SEQ ID NO:2735, SEQ ID NO:2737, SEQ ID NO:2739, SEQ ID NO:nnnn , SEQ ID NO:2743,
 SEQ ID NO:2745, SEQ ID NO:2747, SEQ ID NO:2749, SEQ ID NO:2750, SEQ ID NO:2752,
 SEQ ID NO:2754, SEQ ID NO:2756, SEQ ID NO:2758, SEQ ID NO:2760, SEQ ID NO:2761,
 SEQ ID NO:nnnn , SEQ ID NO:2765, SEQ ID NO:2766, SEQ ID NO:2768, SEQ ID NO:2770,
 15 SEQ ID NO:2772, SEQ ID NO:2774, SEQ ID NO:2776, SEQ ID NO:2778, SEQ ID NO:2780,
 SEQ ID NO:2782, SEQ ID NO:2784, SEQ ID NO:2786, SEQ ID NO:2788, SEQ ID NO:2790,
 SEQ ID NO:2792, SEQ ID NO:2794, SEQ ID NO:2796, SEQ ID NO:2798, SEQ ID NO:2800,
 SEQ ID NO:2802, SEQ ID NO:2804, SEQ ID NO:2806, SEQ ID NO:2808, SEQ ID NO:2810,
 20 SEQ ID NO:2812, SEQ ID NO:2814, SEQ ID NO:2816, SEQ ID NO:2818, SEQ ID NO:2820,
 SEQ ID NO:2822, SEQ ID NO:2824, SEQ ID NO:2826, SEQ ID NO:2828, SEQ ID NO:2830,
 25 SEQ ID NO:2832, SEQ ID NO:2834, SEQ ID NO:2836, SEQ ID NO:2838, SEQ ID NO:2840,
 SEQ ID NO:2842, SEQ ID NO:2844, SEQ ID NO:2846, SEQ ID NO:2848, SEQ ID NO:2850,
 SEQ ID NO:2852, SEQ ID NO:2854, SEQ ID NO:2856, SEQ ID NO:2858, SEQ ID NO:2860,
 SEQ ID NO:2862, SEQ ID NO:2864, SEQ ID NO:2866, SEQ ID NO:2868, SEQ ID NO:2870,
 30 SEQ ID NO:2872, SEQ ID NO:2874, SEQ ID NO:2876, SEQ ID NO:2878, SEQ ID NO:2880,
 SEQ ID NO:2882, SEQ ID NO:2884, SEQ ID NO:2886, SEQ ID NO:2888, SEQ ID NO:2890,
 SEQ ID NO:2892, SEQ ID NO:2894, SEQ ID NO:2896, SEQ ID NO:2898, SEQ ID NO:2900,
 SEQ ID NO:2902, SEQ ID NO:2904, SEQ ID NO:2906, SEQ ID NO:2908, SEQ ID NO:2910,
 35 SEQ ID NO:2912, SEQ ID NO:2914, SEQ ID NO:2916, SEQ ID NO:2918, SEQ ID NO:2920,
 SEQ ID NO:2922, SEQ ID NO:2924, SEQ ID NO:2926, SEQ ID NO:2928, SEQ ID NO:2930,
 SEQ ID NO:2932, SEQ ID NO:2934, SEQ ID NO:2936, SEQ ID NO:2938, SEQ ID NO:2940,
 SEQ ID NO:2942, SEQ ID NO:2944, SEQ ID NO:2946, SEQ ID NO:2948, SEQ ID NO:2950,
 SEQ ID NO:2952, SEQ ID NO:2954, SEQ ID NO:2956, SEQ ID NO:2958, SEQ ID NO:2960,
 40 SEQ ID NO:2962, SEQ ID NO:nnnn , SEQ ID NO:2966, SEQ ID NO:2967, SEQ ID NO:2969,
 SEQ ID NO:2971, SEQ ID NO:2973, SEQ ID NO:2975, SEQ ID NO:2977, SEQ ID NO:2979,
 SEQ ID NO:2981, SEQ ID NO:2983, SEQ ID NO:2985, SEQ ID NO:2987, SEQ ID NO:2989,
 SEQ ID NO:2991, SEQ ID NO:2993, SEQ ID NO:2995, SEQ ID NO:2997, SEQ ID NO:2999,
 SEQ ID NO:3001, SEQ ID NO:3003, SEQ ID NO:3005, SEQ ID NO:3007, SEQ ID NO:3009,
 45 SEQ ID NO:3011, SEQ ID NO:3013, SEQ ID NO:3015, SEQ ID NO:3017, SEQ ID NO:3019,
 SEQ ID NO:3021, SEQ ID NO:3023, SEQ ID NO:3025, SEQ ID NO:3027, SEQ ID NO:3029,
 SEQ ID NO:3031, SEQ ID NO:3033, SEQ ID NO:3035, SEQ ID NO:3037, SEQ ID NO:3039,
 SEQ ID NO:nnnn , SEQ ID NO:nnr.n , SEQ ID NO:3045, SEQ ID NO:3047, SEQ ID NO:3049,
 SEQ ID NO:3051, SEQ ID NO:3053, SEQ ID NO:3055, SEQ ID NO:3057, SEQ ID NO:3059,
 50 SEQ ID NO:3061, SEQ ID NO:3063, SEQ ID NO:3065, SEQ ID NO:3067, SEQ ID NO:3069,
 SEQ ID NO:3071, SEQ ID NO:3073, SEQ ID NO:3075, SEQ ID NO:3077, SEQ ID NO:3079,
 SEQ ID NO:3081, SEQ ID NO:3083, SEQ ID NO:3085, SEQ ID NO:3087, SEQ ID NO:3089,
 SEQ ID NO:3091, SEQ ID NO:3093, SEQ ID NO:3095, SEQ ID NO:3097, SEQ ID NO:3099,
 55 SEQ ID NO:3101, SEQ ID NO:3103, SEQ ID NO:3105, SEQ ID NO:3107, SEQ ID NO:3109,
 SEQ ID NO:3111, SEQ ID NO:3113, SEQ ID NO:3115, SEQ ID NO:3117, SEQ ID NO:3119,
 SEQ ID NO:3121, SEQ ID NO:3123, SEQ ID NO:3125, SEQ ID NO:3127, SEQ ID NO:3129,
 SEQ ID NO:3131, SEQ ID NO:3133, SEQ ID NO:3135, SEQ ID NO:3137, SEQ ID NO:3139,
 SEQ ID NO:3140, SEQ ID NO:3142, SEQ ID NO:3143, SEQ ID NO:3144, SEQ ID NO:3146,
 SEQ ID NO:3148, SEQ ID NO:3150, SEQ ID NO:3152, SEQ ID NO:3154, SEQ ID NO:3156,
 60 SEQ ID NO:3158, SEQ ID NO:nnnn , SEQ ID NO:3162, SEQ ID NO:3164, SEQ ID NO:3166,
 SEQ ID NO:3168, SEQ ID NO:3170, SEQ ID NO:3172, SEQ ID NO:3174, SEQ ID NO:3176,
 SEQ ID NO:3178, SEQ ID NO:3180, SEQ ID NO:3181, SEQ ID NO:3183, SEQ ID NO:3185,
 SEQ ID NO:3187, SEQ ID NO:3188, SEQ ID NO:3189, SEQ ID NO:3191, SEQ ID NO:3193,
 SEQ ID NO:3195, SEQ ID NO:3196, SEQ ID NO:3198, SEQ ID NO:3200, SEQ ID NO:3202,
 65 SEQ ID NO:3204, SEQ ID NO:3206, SEQ ID NO:3208, SEQ ID NO:3210, SEQ ID NO:3212,

EP 1 130 094 A2

5 SEQ ID NO:3214, SEQ ID NO:nnnn , SEQ ID NO:3218, SEQ ID NO:3219, SEQ ID NO:3221,
 SEQ ID NO:3222, SEQ ID NO:3224, SEQ ID NO:3226, SEQ ID NO:nnnn , SEQ ID NO:3230,
 SEQ ID NO:3232, SEQ ID NO:3234, SEQ ID NO:3236, SEQ ID NO:3238, SEQ ID NO:3240,
 10 SEQ ID NO:3242, SEQ ID NO:3244, SEQ ID NO:3246, SEQ ID NO:nnnn , SEQ ID NO:3250,
 SEQ ID NO:3252, SEQ ID NO:3254, SEQ ID NO:3256, SEQ ID NO:3258, SEQ ID NO:3260,
 SEQ ID NO:3262, SEQ ID NO:3264, SEQ ID NO:3266, SEQ ID NO:3268, SEQ ID NO:3270,
 SEQ ID NO:3272, SEQ ID NO:3274, SEQ ID NO:3276, SEQ ID NO:3278, SEQ ID NO:3280,
 15 SEQ ID NO:3282, SEQ ID NO:3284, SEQ ID NO:3286, SEQ ID NO:3288, SEQ ID NO:3290,
 SEQ ID NO:3292, SEQ ID NO:3294, SEQ ID NO:3296, SEQ ID NO:3298, SEQ ID NO:3300,
 SEQ ID NO:3301, SEQ ID NO:3303, SEQ ID NO:3305, SEQ ID NO:3307, SEQ ID NO:3309,
 SEQ ID NO:3311, SEQ ID NO:3312, SEQ ID NO:3314, SEQ ID NO:3316, SEQ ID NO:3318
 SEQ ID NO:3320, SEQ ID NO:3322, SEQ ID NO:3324, SEQ ID NO:3326, SEQ ID NO:3328,
 20 SEQ ID NO:3330, SEQ ID NO:3332, SEQ ID NO:3334, SEQ ID NO:3336, SEQ ID NO:3338,
 SEQ ID NO:nnnn , SEQ ID NO:3342, SEQ ID NO:3344, SEQ ID NO:3346, SEQ ID NO:3348,
 SEQ ID NO:3350, SEQ ID NO:3352, SEQ ID NO:3354, SEQ ID NO:3356, SEQ ID NO:3358,
 SEQ ID NO:3360, SEQ ID NO:3362, SEQ ID NO:3364, SEQ ID NO:3366, SEQ ID NO:3368,
 SEQ ID NO:3370, SEQ ID NO:3372, SEQ ID NO:3374, SEQ ID NO:3376, SEQ ID NO:3378,
 25 SEQ ID NO:3380, SEQ ID NO:3382, SEQ ID NO:3384, SEQ ID NO:3386, SEQ ID NO:3388,
 SEQ ID NO:3390, SEQ ID NO:3392, SEQ ID NO:3394, SEQ ID NO:3396, SEQ ID NO:3398,
 SEQ ID NO:3400, SEQ ID NO:3402, SEQ ID NO:3404, SEQ ID NO:nnnn , SEQ ID NO:3408,
 SEQ ID NO:3410, SEQ ID NO:3411, SEQ ID NO:3413, SEQ ID NO:3415, SEQ ID NO:3417,
 SEQ ID NO:3419, SEQ ID NO:nnnn , SEQ ID NO:3423, SEQ ID NO:3425, SEQ ID NO:3427,
 30 SEQ ID NO:3429, SEQ ID NO:3431, SEQ ID NO:3433, SEQ ID NO:3435, SEQ ID NO:3437,
 SEQ ID NO:3439, SEQ ID NO:3441, SEQ ID NO:3443, SEQ ID NO:3445, SEQ ID NO:3447,
 SEQ ID NO:3449, SEQ ID NO:3451, SEQ ID NO:3453, SEQ ID NO:3455, SEQ ID NO:3457,
 SEQ ID NO:3459, SEQ ID NO:3461, SEQ ID NO:3463, SEQ ID NO:3465, SEQ ID NO:3467,
 SEQ ID NO:3469, SEQ ID NO:3471, SEQ ID NO:3473, SEQ ID NO:3475, SEQ ID NO:3477,
 35 SEQ ID NO:3479, SEQ ID NO:3481, SEQ ID NO:3482, SEQ ID NO:3484, SEQ ID NO:3486,
 SEQ ID NO:3488, SEQ ID NO:3490, SEQ ID NO:3492, SEQ ID NO:3494, SEQ ID NO:3496,
 SEQ ID NO:3498, SEQ ID NO:3500, SEQ ID NO:3502, SEQ ID NO:3504, SEQ ID NO:3506,
 SEQ ID NO:3508, SEQ ID NO:3509, SEQ ID NO:3511, SEQ ID NO:3513, SEQ ID NO:3515
 SEQ ID NO:3517, SEQ ID NO:3519, SEQ ID NO:3521, SEQ ID NO:3523, SEQ ID NO:3525,
 40 SEQ ID NO:3527, SEQ ID NO:3529, SEQ ID NO:3531, SEQ ID NO:3533, SEQ ID NO:3535,
 SEQ ID NO:3537, SEQ ID NO:3539, SEQ ID NO:3541, SEQ ID NO:3543, SEQ ID NO:3545,
 SEQ ID NO:3547, SEQ ID NO:3549, SEQ ID NO:3551, SEQ ID NO:3553, SEQ ID NO:3555,
 SEQ ID NO:3557, SEQ ID NO:3559, SEQ ID NO:3561, SEQ ID NO:3563, SEQ ID NO:nnnn ,
 SEQ ID NO:3567, SEQ ID NO:3569, SEQ ID NO:3571, SEQ ID NO:3573, SEQ ID NO:3575,
 45 SEQ ID NO:3577, SEQ ID NO:3579, SEQ ID NO:3581, SEQ ID NO:3583, SEQ ID NO:3585,
 SEQ ID NO:3587, SEQ ID NO:3589 , SEQ ID NO:3595, SEQ ID NO:3597, SEQ ID NO:3599,
 SEQ ID NO:3601, SEQ ID NO:3603, SEQ ID NO:3605, SEQ ID NO:3607, SEQ ID NO:3609,
 SEQ ID NO:3610, SEQ ID NO:3611, SEQ ID NO:3612, SEQ ID NO:3613, SEQ ID NO:3615,
 SEQ ID NO:3617, SEQ ID NO:3619, SEQ ID NO:3621, SEQ ID NO:3622, SEQ ID NO:3624,
 50 SEQ ID NO:3626, SEQ ID NO:3627, SEQ ID NO:3628, SEQ ID NO:3629, SEQ ID NO:3631,
 SEQ ID NO:3633, SEQ ID NO:3634, SEQ ID NO:3636, SEQ ID NO:3638, SEQ ID NO:3640,
 SEQ ID NO:3642, SEQ ID NO:3644, SEQ ID NO:3645, SEQ ID NO:3647, SEQ ID NO:3649,
 SEQ ID NO:3651, SEQ ID NO:3653, SEQ ID NO:3655, SEQ ID NO:3657, SEQ ID NO:3658,
 SEQ ID NO:3660, SEQ ID NO:3662, SEQ ID NO:3664, SEQ ID NO:3666, SEQ ID NO:3668,
 55 SEQ ID NO:3670, SEQ ID NO:3672, SEQ ID NO:3674, SEQ ID NO:3676, SEQ ID NO:3678,
 SEQ ID NO:3680, SEQ ID NO:3682, SEQ ID NO:3683, SEQ ID NO:3685, SEQ ID NO:3686,
 SEQ ID NO:3688, SEQ ID NO:3690, SEQ ID NO:3692, SEQ ID NO:3694, SEQ ID NO:3696,
 SEQ ID NO:3698, SEQ ID NO:3700, SEQ ID NO:3702, SEQ ID NO:3704, SEQ ID NO:3706,
 SEQ ID NO:3708, SEQ ID NO:3710, SEQ ID NO:3712, SEQ ID NO:3714, SEQ ID NO:3716,

SEQ ID NO:3718, SEQ ID NO:3720, SEQ ID NO:3722, SEQ ID NO:3724, SEQ ID NO:3726,
 SEQ ID NO:3728, SEQ ID NO:3730, SEQ ID NO:3732, SEQ ID NO:3734, SEQ ID NO:3736,
 SEQ ID NO:3738, SEQ ID NO:3740, SEQ ID NO:3741, SEQ ID NO:3743, SEQ ID NO:3745,
 5 SEQ ID NO:3747, SEQ ID NO:3749, SEQ ID NO:3751, SEQ ID NO:3753, SEQ ID NO:3755,
 SEQ ID NO:3756, SEQ ID NO:3758, SEQ ID NO:3760, SEQ ID NO:3761, SEQ ID NO:3763,
 SEQ ID NO:3765, SEQ ID NO:3767, SEQ ID NO:3769, SEQ ID NO:3771, SEQ ID NO:3773,
 SEQ ID NO:3775, SEQ ID NO:3777
 10 SEQ ID NO:3779, SEQ ID NO:3781, SEQ ID NO:3783, SEQ ID NO:3785, SEQ ID NO:3787,
 SEQ ID NO:3789, SEQ ID NO:3791, SEQ ID NO:3793, SEQ ID NO:3795, SEQ ID NO:3797,
 SEQ ID NO:3799, SEQ ID NO:3801, SEQ ID NO:3803, SEQ ID NO:3805, SEQ ID NO:3806,
 SEQ ID NO:3808, SEQ ID NO:3810, SEQ ID NO:3812, SEQ ID NO:3814, SEQ ID NO:3816,
 15 SEQ ID NO:3818, SEQ ID NO:3820, SEQ ID NO:3822, SEQ ID NO:3824, SEQ ID NO:3826,
 SEQ ID NO:3828, SEQ ID NO:3830, SEQ ID NO:3832, SEQ ID NO:3834, SEQ ID NO:3836,
 SEQ ID NO:3838, SEQ ID NO:3840, SEQ ID NO:3842, SEQ ID NO:3844, SEQ ID NO:3846,
 SEQ ID NO:3848, SEQ ID NO:3852, SEQ ID NO:3854, SEQ ID NO:3856, SEQ ID NO:3858,
 SEQ ID NO:3860, SEQ ID NO:3862, SEQ ID NO:3864, SEQ ID NO:3866, SEQ ID NO:3868,
 20 SEQ ID NO:3870, SEQ ID NO:3872, SEQ ID NO:3874, SEQ ID NO:3876, SEQ ID NO:3878,
 SEQ ID NO:3880, SEQ ID NO:3882, SEQ ID NO:3884, SEQ ID NO:3886, SEQ ID NO:3888,
 SEQ ID NO:3890, SEQ ID NO:3892, SEQ ID NO:3894, SEQ ID NO:3896, SEQ ID NO:3898,
 SEQ ID NO:3900, SEQ ID NO:3902, SEQ ID NO:3904, SEQ ID NO:3906, SEQ ID NO:3908,
 25 SEQ ID NO:3910, SEQ ID NO:3912, SEQ ID NO:3914, SEQ ID NO:3916, SEQ ID NO:3918,
 SEQ ID NO:3920, SEQ ID NO:3922, SEQ ID NO:3924, SEQ ID NO:3926, SEQ ID NO:3928,
 SEQ ID NO:3930, SEQ ID NO:3932, SEQ ID NO:3934, SEQ ID NO:3936, SEQ ID NO:3938,
 SEQ ID NO:3940, SEQ ID NO:3942, SEQ ID NO:3944, SEQ ID NO:3946, SEQ ID NO:3948,
 SEQ ID NO:3950, SEQ ID NO:3952, SEQ ID NO:3954, SEQ ID NO:3956, SEQ ID NO:3958,
 30 SEQ ID NO:3960, SEQ ID NO:3962, SEQ ID NO:3964, SEQ ID NO:3965, SEQ ID NO:3966,
 SEQ ID NO:3968, SEQ ID NO:3970, SEQ ID NO:3972, SEQ ID NO:3974, SEQ ID NO:3976,
 SEQ ID NO:3977, SEQ ID NO:3979, SEQ ID NO:3981, SEQ ID NO:3983, SEQ ID NO:3985,
 SEQ ID NO:3987, SEQ ID NO:3988, SEQ ID NO:3990, SEQ ID NO:3992, SEQ ID NO:3994,
 SEQ ID NO:3996, SEQ ID NO:3998, SEQ ID NO:4000, SEQ ID NO:4001, SEQ ID NO:4003,
 35 SEQ ID NO:4004, SEQ ID NO:4006, SEQ ID NO:4008, SEQ ID NO:4010, SEQ ID NO:4012,
 SEQ ID NO:4016, SEQ ID NO:4018, SEQ ID NO:4020, SEQ ID NO:4022, SEQ ID NO:4024,
 SEQ ID NO:4026, SEQ ID NO:4028, SEQ ID NO:4030, SEQ ID NO:4032, SEQ ID NO:4034,
 SEQ ID NO:4036, SEQ ID NO:4038, SEQ ID NO:4040, SEQ ID NO:4042, SEQ ID NO:4044,
 40 SEQ ID NO:4046, SEQ ID NO:4048, SEQ ID NO:4050, SEQ ID NO:4052, SEQ ID NO:4054,
 SEQ ID NO:4056, SEQ ID NO:4058, SEQ ID NO:4060, SEQ ID NO:4062, SEQ ID NO:4064,
 SEQ ID NO:4066, SEQ ID NO:4068, SEQ ID NO:4070, SEQ ID NO:4072, SEQ ID NO:4074,
 SEQ ID NO:4076
 SEQ ID NO:4078, SEQ ID NO:4080, SEQ ID NO:4082, SEQ ID NO:4084, SEQ ID NO:4086,
 45 SEQ ID NO:4088, SEQ ID NO:4090, SEQ ID NO:4092, SEQ ID NO:4094, SEQ ID NO:4096,
 SEQ ID NO:4098, SEQ ID NO:4099, SEQ ID NO:4101, SEQ ID NO:4103, SEQ ID NO:4105,
 SEQ ID NO:4107, SEQ ID NO:4109, SEQ ID NO:4111, SEQ ID NO:4113, SEQ ID NO:4115,
 SEQ ID NO:4117, SEQ ID NO:4119, SEQ ID NO:4121, SEQ ID NO:4123, SEQ ID NO:4125,
 50 SEQ ID NO:4127, SEQ ID NO:4129, SEQ ID NO:4131, SEQ ID NO:4132, SEQ ID NO:4134,
 SEQ ID NO:4136, SEQ ID NO:4137, SEQ ID NO:4139, SEQ ID NO:4140, SEQ ID NO:4142,
 SEQ ID NO:4144, SEQ ID NO:4146, SEQ ID NO:4148, SEQ ID NO:4150, SEQ ID NO:4152,
 SEQ ID NO:4154, SEQ ID NO:4155, SEQ ID NO:4157, SEQ ID NO:4159, SEQ ID NO:4161,
 SEQ ID NO:4163, SEQ ID NO:4165
 55 SEQ ID NO:4167, SEQ ID NO:4168, SEQ ID NO:4170, SEQ ID NO:4172, SEQ ID NO:4174,
 SEQ ID NO:4176, and SEQ ID NO:4178

(b) a polynucleotide comprising a nucleotide sequence encoding a protein comprising the amino acid sequence

EP 1 130 094 A2

set forth in any one of the following SEQ ID NOs:

5

10

15

20

25

30

35

40

45

50

55

EP 1 130 094 A2

5 SEQ ID NO:2548, SEQ ID NO:nnnn , SEQ ID NO:2552, SEQ ID NO:2554, SEQ ID NO:2556,
 SEQ ID NO:2559, SEQ ID NO:2561, SEQ ID NO:2563, SEQ ID NO:2565, SEQ ID NO:2567,
 SEQ ID NO:2569, SEQ ID NO:2571, SEQ ID NO:2573, SEQ ID NO:2575, SEQ ID NO:2577,
 SEQ ID NO:2579, SEQ ID NO:2582, SEQ ID NO:2585, SEQ ID NO:nnnn , SEQ ID NO:2589,
 SEQ ID NO:2591, SEQ ID NO:2593, SEQ ID NO:2595, SEQ ID NO:2597, SEQ ID NO:2599,
 SEQ ID NO:nnnn , SEQ ID NO:2603, SEQ ID NO:2605, SEQ ID NO:2608, SEQ ID NO:2610,
 SEQ ID NO:2612, SEQ ID NO:2614, SEQ ID NO:2616, SEQ ID NO:2618, SEQ ID NO:2620,
 10 SEQ ID NO:2622, SEQ ID NO:2624, SEQ ID NO:2626, SEQ ID NO:2628, SEQ ID NO:2631,
 SEQ ID NO:2633, SEQ ID NO:2635, SEQ ID NO:2637, SEQ ID NO:2639, SEQ ID NO:2641,
 SEQ ID NO:2643, SEQ ID NO:2645, SEQ ID NO:2647, SEQ ID NO:2649, SEQ ID NO:2651,
 SEQ ID NO:2653, SEQ ID NO:2655, SEQ ID NO:2657, SEQ ID NO:2659, SEQ ID NO:2661,
 SEQ ID NO:2663, SEQ ID NO:2665, SEQ ID NO:2667, SEQ ID NO:2669, SEQ ID NO:2671,
 15 SEQ ID NO:2673, SEQ ID NO:2675, SEQ ID NO:2677, SEQ ID NO:2679, SEQ ID NO:2682,
 SEQ ID NO:2684, SEQ ID NO:2687, SEQ ID NO:2689, SEQ ID NO:2691, SEQ ID NO:2693,
 SEQ ID NO:2695, SEQ ID NO:2697, SEQ ID NO:2699, SEQ ID NO:2701, SEQ ID NO:2703,
 SEQ ID NO:2705, SEQ ID NO:2707, SEQ ID NO:2709, SEQ ID NO:2711, SEQ ID NO:2713,
 SEQ ID NO:2716, SEQ ID NO:2720, SEQ ID NO:2723, SEQ ID NO:2725, SEQ ID NO:2727,
 20 SEQ ID NO:2729, SEQ ID NO:2731, SEQ ID NO:2733, SEQ ID NO:2736, SEQ ID NO:2738,
 SEQ ID NO:2740, SEQ ID NO:nnnn , SEQ ID NO:2744, SEQ ID NO:2746, SEQ ID NO:2748,
 SEQ ID NO:2751, SEQ ID NO:2753, SEQ ID NO:2755, SEQ ID NO:2757, SEQ ID NO:2759
 SEQ ID NO:2762, SEQ ID NO:nnnn , SEQ ID NO:2767, SEQ ID NO:2769, SEQ ID NO:2771,
 SEQ ID NO:2773, SEQ ID NO:2775, SEQ ID NO:2777, SEQ ID NO:2779, SEQ ID NO:2781,
 25 SEQ ID NO:2783, SEQ ID NO:2785, SEQ ID NO:2787, SEQ ID NO:2789, SEQ ID NO:2791,
 SEQ ID NO:2793, SEQ ID NO:2795, SEQ ID NO:2797, SEQ ID NO:2799, SEQ ID NO:2801,
 SEQ ID NO:2803, SEQ ID NO:2805, SEQ ID NO:2807, SEQ ID NO:2809, SEQ ID NO:2811,
 SEQ ID NO:2813, SEQ ID NO:2815, SEQ ID NO:2817, SEQ ID NO:2819, SEQ ID NO:2821,
 SEQ ID NO:2823, SEQ ID NO:2825, SEQ ID NO:2827, SEQ ID NO:2829, SEQ ID NO:2831,
 30 SEQ ID NO:2833, SEQ ID NO:2835, SEQ ID NO:2837, SEQ ID NO:2839, SEQ ID NO:2841,
 SEQ ID NO:2843, SEQ ID NO:2845, SEQ ID NO:2847, SEQ ID NO:2849, SEQ ID NO:2851,
 SEQ ID NO:2853, SEQ ID NO:2855, SEQ ID NO:2857, SEQ ID NO:2859, SEQ ID NO:2861,
 SEQ ID NO:2863, SEQ ID NO:2865, SEQ ID NO:2867, SEQ ID NO:2869, SEQ ID NO:2871,
 SEQ ID NO:2873, SEQ ID NO:2875, SEQ ID NO:2877, SEQ ID NO:2879, SEQ ID NO:2881,
 35 SEQ ID NO:2883, SEQ ID NO:2885, SEQ ID NO:2887, SEQ ID NO:2889, SEQ ID NO:2891,
 SEQ ID NO:2893, SEQ ID NO:2895, SEQ ID NO:2897, SEQ ID NO:2899, SEQ ID NO:2901,
 SEQ ID NO:2903, SEQ ID NO:2905, SEQ ID NO:2907, SEQ ID NO:2909, SEQ ID NO:2911,
 SEQ ID NO:2913, SEQ ID NO:2915, SEQ ID NO:2917, SEQ ID NO:2919, SEQ ID NO:2921,
 SEQ ID NO:2923, SEQ ID NO:2925, SEQ ID NO:2927, SEQ ID NO:2929, SEQ ID NO:2931,
 40 SEQ ID NO:2933, SEQ ID NO:2935, SEQ ID NO:2937, SEQ ID NO:2939, SEQ ID NO:2941,
 SEQ ID NO:2943, SEQ ID NO:2945, SEQ ID NO:2947, SEQ ID NO:2949, SEQ ID NO:2951,
 SEQ ID NO:2953, SEQ ID NO:2955, SEQ ID NO:2957, SEQ ID NO:2959, SEQ ID NO:2961
 SEQ ID NO:2963, SEQ ID NO:nnnn , SEQ ID NO:2968, SEQ ID NO:2970, SEQ ID NO:2972,
 45 SEQ ID NO:2974, SEQ ID NO:2976, SEQ ID NO:2978, SEQ ID NO:2980, SEQ ID NO:2982,
 SEQ ID NO:2984, SEQ ID NO:2986, SEQ ID NO:2988, SEQ ID NO:2990, SEQ ID NO:2992,
 SEQ ID NO:2994, SEQ ID NO:2996, SEQ ID NO:2998, SEQ ID NO:3000, SEQ ID NO:3002,
 SEQ ID NO:3004, SEQ ID NO:3006, SEQ ID NO:3008, SEQ ID NO:3010, SEQ ID NO:3012,
 SEQ ID NO:3014, SEQ ID NO:3016, SEQ ID NO:3018, SEQ ID NO:3020, SEQ ID NO:3022,
 50 SEQ ID NO:3024, SEQ ID NO:3026, SEQ ID NO:3028, SEQ ID NO:3030, SEQ ID NO:3032,
 SEQ ID NO:3034, SEQ ID NO:3036, SEQ ID NO:3038, SEQ ID NO:3040, SEQ ID NO:nnnn ,
 SEQ ID NO:nnnn , SEQ ID NO:3046, SEQ ID NO:3048, SEQ ID NO:3050, SEQ ID NO:3052,
 SEQ ID NO:3054, SEQ ID NO:3056, SEQ ID NO:3058, SEQ ID NO:3060, SEQ ID NO:3062,
 SEQ ID NO:3064, SEQ ID NO:3066, SEQ ID NO:3068, SEQ ID NO:3070, SEQ ID NO:3072,
 55 SEQ ID NO:3074, SEQ ID NO:3076, SEQ ID NO:3078, SEQ ID NO:3080, SEQ ID NO:3082.

EP 1 130 094 A2

5 SEQ ID NO:3084, SEQ ID NO:3086, SEQ ID NO:3088, SEQ ID NO:3090, SEQ ID NO:3092,
 SEQ ID NO:3094, SEQ ID NO:3096, SEQ ID NO:3098, SEQ ID NO:3100, SEQ ID NO:3102,
 SEQ ID NO:3104, SEQ ID NO:3106, SEQ ID NO:3108, SEQ ID NO:3110, SEQ ID NO:3112,
 SEQ ID NO:3114, SEQ ID NO:3116, SEQ ID NO:3118, SEQ ID NO:3120, SEQ ID NO:3122,
 SEQ ID NO:3124, SEQ ID NO:3126, SEQ ID NO:3128, SEQ ID NO:3130, SEQ ID NO:3132,
 SEQ ID NO:3134, SEQ ID NO:3136, SEQ ID NO:3138, SEQ ID NO:3141, SEQ ID NO:3145,
 SEQ ID NO:3147, SEQ ID NO:3149, SEQ ID NO:3151, SEQ ID NO:3153, SEQ ID NO:3155,
 10 SEQ ID NO:3157, SEQ ID NO:3159, SEQ ID NO:nnnn , SEQ ID NO:3163, SEQ ID NO:3165
 SEQ ID NO:3167, SEQ ID NO:3169, SEQ ID NO:3171, SEQ ID NO:3173, SEQ ID NO:3175,
 SEQ ID NO:3177, SEQ ID NO:3179, SEQ ID NO:3182, SEQ ID NO:3184, SEQ ID NO:3186,
 SEQ ID NO:3190, SEQ ID NO:3192, SEQ ID NO:3194, SEQ ID NO:3197, SEQ ID NO:3199,
 SEQ ID NO:3201, SEQ ID NO:3203, SEQ ID NO:3205, SEQ ID NO:3207, SEQ ID NO:3209,
 15 SEQ ID NO:3211, SEQ ID NO:3213, SEQ ID NO:3215, SEQ ID NO:nnnn , SEQ ID NO:3220,
 SEQ ID NO:3223, SEQ ID NO:3225, SEQ ID NO:3227, SEQ ID NO:nnnn , SEQ ID NO:3231,
 SEQ ID NO:3233, SEQ ID NO:3235, SEQ ID NO:3237, SEQ ID NO:3239, SEQ ID NO:3241,
 SEQ ID NO:3243, SEQ ID NO:3245, SEQ ID NO:3247, SEQ ID NO:nnnn , SEQ ID NO:3251,
 SEQ ID NO:3253, SEQ ID NO:3255, SEQ ID NO:3257, SEQ ID NO:3259, SEQ ID NO:3261,
 20 SEQ ID NO:3263, SEQ ID NO:3265, SEQ ID NO:3267, SEQ ID NO:3269, SEQ ID NO:3271,
 SEQ ID NO:3273, SEQ ID NO:3275, SEQ ID NO:3277, SEQ ID NO:3279, SEQ ID NO:3281,
 SEQ ID NO:3283, SEQ ID NO:3285, SEQ ID NO:3287, SEQ ID NO:3289, SEQ ID NO:3291,
 SEQ ID NO:3293, SEQ ID NO:3295, SEQ ID NO:3297, SEQ ID NO:3299, SEQ ID NO:3302,
 SEQ ID NO:3304, SEQ ID NO:3306, SEQ ID NO:3308, SEQ ID NO:3310, SEQ ID NO:3313,
 25 SEQ ID NO:3315, SEQ ID NO:3317, SEQ ID NO:3319, SEQ ID NO:3321, SEQ ID NO:3323,
 SEQ ID NO:3325, SEQ ID NO:3327, SEQ ID NO:3329, SEQ ID NO:3331, SEQ ID NO:3333,
 SEQ ID NO:3335, SEQ ID NO:3337, SEQ ID NO:3339, SEQ ID NO:nnnn , SEQ ID NO:3343,
 SEQ ID NO:3345, SEQ ID NO:3347, SEQ ID NO:3349, SEQ ID NO:3351, SEQ ID NO:3353,
 SEQ ID NO:3355, SEQ ID NO:3357, SEQ ID NO:3359, SEQ ID NO:3361, SEQ ID NO:3363,
 30 SEQ ID NO:3365, SEQ ID NO:3367, SEQ ID NO:3369, SEQ ID NO:3371, SEQ ID NO:3373
 SEQ ID NO:3375, SEQ ID NO:3377, SEQ ID NO:3379, SEQ ID NO:3381, SEQ ID NO:3383,
 SEQ ID NO:3385, SEQ ID NO:3387, SEQ ID NO:3389, SEQ ID NO:3391, SEQ ID NO:3393,
 SEQ ID NO:3395, SEQ ID NO:3397, SEQ ID NO:3399, SEQ ID NO:3401, SEQ ID NO:3403,
 SEQ ID NO:3405, SEQ ID NO:nnnn , SEQ ID NO:3409, SEQ ID NO:3412, SEQ ID NO:3414,
 35 SEQ ID NO:3416, SEQ ID NO:3418, SEQ ID NO:3420, SEQ ID NO:nnnn , SEQ ID NO:3424,
 SEQ ID NO:3426, SEQ ID NO:3428, SEQ ID NO:3430, SEQ ID NO:3432, SEQ ID NO:3434,
 SEQ ID NO:3436, SEQ ID NO:3438, SEQ ID NO:3440, SEQ ID NO:3442, SEQ ID NO:3444,
 SEQ ID NO:3446, SEQ ID NO:3448, SEQ ID NO:3450, SEQ ID NO:3452, SEQ ID NO:3454,
 SEQ ID NO:3456, SEQ ID NO:3458, SEQ ID NO:3460, SEQ ID NO:3462, SEQ ID NO:3464,
 40 SEQ ID NO:3466, SEQ ID NO:3468, SEQ ID NO:3470, SEQ ID NO:3472, SEQ ID NO:3474,
 SEQ ID NO:3476, SEQ ID NO:3478, SEQ ID NO:3480, SEQ ID NO:3483, SEQ ID NO:3485,
 SEQ ID NO:3487, SEQ ID NO:3489, SEQ ID NO:3491, SEQ ID NO:3493, SEQ ID NO:3495,
 SEQ ID NO:3497, SEQ ID NO:3499, SEQ ID NO:3501, SEQ ID NO:3503, SEQ ID NO:3505,
 45 SEQ ID NO:3507, SEQ ID NO:3510, SEQ ID NO:3512, SEQ ID NO:3514, SEQ ID NO:3516,
 SEQ ID NO:3518, SEQ ID NO:3520, SEQ ID NO:3522, SEQ ID NO:3524, SEQ ID NO:3526,
 SEQ ID NO:3528, SEQ ID NO:3530, SEQ ID NO:3532, SEQ ID NO:3534, SEQ ID NO:3536,
 SEQ ID NO:3538, SEQ ID NO:3540, SEQ ID NO:3542, SEQ ID NO:3544, SEQ ID NO:3546,
 SEQ ID NO:3548, SEQ ID NO:3550, SEQ ID NO:3552, SEQ ID NO:3554, SEQ ID NO:3556,
 50 SEQ ID NO:3558, SEQ ID NO:3560, SEQ ID NO:3562, SEQ ID NO:3564, SEQ ID NO:nnnn ,
 SEQ ID NO:3568, SEQ ID NO:3570, SEQ ID NO:3572, SEQ ID NO:3574, SEQ ID NO:3576
 SEQ ID NO:3578, SEQ ID NO:3580, SEQ ID NO:3582, SEQ ID NO:3584, SEQ ID NO:3586,
 SEQ ID NO:3588, SEQ ID NO:3590
 SEQ ID NO:3596, SEQ ID NO:3598, SEQ ID NO:3600, SEQ ID NO:3602, SEQ ID NO:3604,
 55 SEQ ID NO:3606, SEQ ID NO:3608, SEQ ID NO:3614, SEQ ID NO:3616, SEQ ID NO:3618,

EP 1 130 094 A2

5 SEQ ID NO:3620, SEQ ID NO:3623, SEQ ID NO:3625, SEQ ID NO:3630, SEQ ID NO:3632,
 SEQ ID NO:3635, SEQ ID NO:3637, SEQ ID NO:3639, SEQ ID NO:3641, SEQ ID NO:3643,
 SEQ ID NO:3646, SEQ ID NO:3648, SEQ ID NO:3650, SEQ ID NO:3652, SEQ ID NO:3654,
 SEQ ID NO:3656, SEQ ID NO:3659, SEQ ID NO:3661, SEQ ID NO:3663, SEQ ID NO:3665,
 SEQ ID NO:3667, SEQ ID NO:3669, SEQ ID NO:3671, SEQ ID NO:3673, SEQ ID NO:3675,
 SEQ ID NO:3677, SEQ ID NO:3679, SEQ ID NO:3681, SEQ ID NO:3684, SEQ ID NO:3687,
 10 SEQ ID NO:3689, SEQ ID NO:3691, SEQ ID NO:3693, SEQ ID NO:3695, SEQ ID NO:3697,
 SEQ ID NO:3699, SEQ ID NO:3701, SEQ ID NO:3703, SEQ ID NO:3705, SEQ ID NO:3707,
 SEQ ID NO:3709, SEQ ID NO:3711, SEQ ID NO:3713, SEQ ID NO:3715, SEQ ID NO:3717,
 SEQ ID NO:3719, SEQ ID NO:3721, SEQ ID NO:3723, SEQ ID NO:3725, SEQ ID NO:3727,
 SEQ ID NO:3729, SEQ ID NO:3731, SEQ ID NO:3733, SEQ ID NO:3735, SEQ ID NO:3737,
 SEQ ID NO:3739, SEQ ID NO:3742, SEQ ID NO:3744, SEQ ID NO:3746, SEQ ID NO:3748,
 15 SEQ ID NO:3750, SEQ ID NO:3752, SEQ ID NO:3754, SEQ ID NO:3757, SEQ ID NO:3759,
 SEQ ID NO:3762, SEQ ID NO:3764, SEQ ID NO:3766, SEQ ID NO:3768, SEQ ID NO:3770,
 SEQ ID NO:3772, SEQ ID NO:3774, SEQ ID NO:3776, SEQ ID NO:3778, SEQ ID NO:3780,
 SEQ ID NO:3782, SEQ ID NO:3784, SEQ ID NO:3786, SEQ ID NO:3788, SEQ ID NO:3790,
 SEQ ID NO:3792, SEQ ID NO:3794, SEQ ID NO:3796, SEQ ID NO:3798, SEQ ID NO:3800,
 20 SEQ ID NO:3802, SEQ ID NO:3804, SEQ ID NO:3807, SEQ ID NO:3809, SEQ ID NO:3811
 SEQ ID NO:3813, SEQ ID NO:3815, SEQ ID NO:3817, SEQ ID NO:3819, SEQ ID NO:3821,
 SEQ ID NO:3823, SEQ ID NO:3825, SEQ ID NO:3827, SEQ ID NO:3829, SEQ ID NO:3831,
 SEQ ID NO:3833, SEQ ID NO:3835, SEQ ID NO:3837, SEQ ID NO:3839, SEQ ID NO:3841,
 SEQ ID NO:3843, SEQ ID NO:3845, SEQ ID NO:3847, SEQ ID NO:3849, SEQ ID NO:3853,
 25 SEQ ID NO:3855, SEQ ID NO:3857, SEQ ID NO:3859, SEQ ID NO:3861, SEQ ID NO:3863,
 SEQ ID NO:3865, SEQ ID NO:3867, SEQ ID NO:3869, SEQ ID NO:3871, SEQ ID NO:3873,
 SEQ ID NO:3875, SEQ ID NO:3877, SEQ ID NO:3879, SEQ ID NO:3881, SEQ ID NO:3883,
 SEQ ID NO:3885, SEQ ID NO:3887, SEQ ID NO:3889, SEQ ID NO:3891, SEQ ID NO:3893,
 SEQ ID NO:3895, SEQ ID NO:3897, SEQ ID NO:3899, SEQ ID NO:3901, SEQ ID NO:3903,
 30 SEQ ID NO:3905, SEQ ID NO:3907, SEQ ID NO:3909, SEQ ID NO:3911, SEQ ID NO:3913,
 SEQ ID NO:3915, SEQ ID NO:3917, SEQ ID NO:3919, SEQ ID NO:3921, SEQ ID NO:3923,
 SEQ ID NO:3925, SEQ ID NO:3927, SEQ ID NO:3929, SEQ ID NO:3931, SEQ ID NO:3933,
 SEQ ID NO:3935, SEQ ID NO:3937, SEQ ID NO:3939, SEQ ID NO:3941, SEQ ID NO:3943,
 SEQ ID NO:3945, SEQ ID NO:3947, SEQ ID NO:3949, SEQ ID NO:3951, SEQ ID NO:3953,
 35 SEQ ID NO:3955, SEQ ID NO:3957, SEQ ID NO:3959, SEQ ID NO:3961, SEQ ID NO:3963,
 SEQ ID NO:3967, SEQ ID NO:3969, SEQ ID NO:3971, SEQ ID NO:3973, SEQ ID NO:3975,
 SEQ ID NO:3978, SEQ ID NO:3980, SEQ ID NO:3982, SEQ ID NO:3984, SEQ ID NO:3986,
 SEQ ID NO:3989, SEQ ID NO:3991, SEQ ID NO:3993, SEQ ID NO:3995, SEQ ID NO:3997,
 SEQ ID NO:3999, SEQ ID NO:4002, SEQ ID NO:4005, SEQ ID NO:4007, SEQ ID NO:4009,
 40 SEQ ID NO:4011, SEQ ID NO:4013, SEQ ID NO:4017, SEQ ID NO:4019, SEQ ID NO:4021
 SEQ ID NO:4023, SEQ ID NO:4025, SEQ ID NO:4027, SEQ ID NO:4029, SEQ ID NO:4031,
 SEQ ID NO:4033, SEQ ID NO:4035, SEQ ID NO:4037, SEQ ID NO:4039, SEQ ID NO:4041,
 SEQ ID NO:4043, SEQ ID NO:4045, SEQ ID NO:4047, SEQ ID NO:4049, SEQ ID NO:4051,
 45 SEQ ID NO:4053, SEQ ID NO:4055, SEQ ID NO:4057, SEQ ID NO:4059, SEQ ID NO:4061,
 SEQ ID NO:4063, SEQ ID NO:4065, SEQ ID NO:4067, SEQ ID NO:4069, SEQ ID NO:4071,
 SEQ ID NO:4073, SEQ ID NO:4075, SEQ ID NO:4077
 SEQ ID NO:4079, SEQ ID NO:4081, SEQ ID NO:4083, SEQ ID NO:4085, SEQ ID NO:4087,
 SEQ ID NO:4089, SEQ ID NO:4091, SEQ ID NO:4093, SEQ ID NO:4095, SEQ ID NO:4097,
 50 SEQ ID NO:4100, SEQ ID NO:4102, SEQ ID NO:4104, SEQ ID NO:4106, SEQ ID NO:4108,
 SEQ ID NO:4110, SEQ ID NO:4112, SEQ ID NO:4114, SEQ ID NO:4116, SEQ ID NO:4118,
 SEQ ID NO:4120, SEQ ID NO:4122, SEQ ID NO:4124, SEQ ID NO:4126, SEQ ID NO:4128,
 SEQ ID NO:4130, SEQ ID NO:4133, SEQ ID NO:4135, SEQ ID NO:4138, SEQ ID NO:4141,
 SEQ ID NO:4143, SEQ ID NO:4145, SEQ ID NO:4147, SEQ ID NO:4149, SEQ ID NO:4151,
 55 SEQ ID NO:4153, SEQ ID NO:4156, SEQ ID NO:4158, SEQ ID NO:4160, SEQ ID NO:4162,

SEQ ID NO:4164, SEQ ID NO:4166
 SEQ ID NO:4169, SEQ ID NO:4171, SEQ ID NO:4173, SEQ ID NO:4175, SEQ ID NO:4177,
 and SEQ ID NO:4179

5

(c) a polynucleotide comprising a nucleotide sequence encoding a protein comprising an amino acid sequence selected from the amino acid sequences of (b), in which one or more amino acids are substituted, deleted, inserted, and/or added, wherein said protein is functionally equivalent to the protein comprising said amino acid sequence selected from the amino acid sequences of (b);

10

(d) a polynucleotide that hybridizes with a polynucleotide comprising a nucleotide sequence selected from the nucleotide sequences of (a), and that comprises a nucleotide sequence encoding a protein functionally equivalent to the protein encoded by the nucleotide sequence selected from the nucleotide sequences of (a);

15

(e) a polynucleotide comprising a nucleotide sequence encoding a partial amino acid sequence of a protein encoded by the polynucleotide of (a) to (d);

(f) a polynucleotide comprising a nucleotide sequence with at least 70% identity to the nucleotide sequence of (a).

20

9. A substantially pure protein encoded by the polynucleotide of claim 8.

10. An antibody against the protein or peptide of any one of claims 6, 7, and 9.

11. A vector comprising the polynucleotide of claim 5 or 8.

25

12. A transformant carrying the polynucleotide of claim 5 or 8, or the vector of claim 11.

13. A transformant expressively carrying the polynucleotide of claim 5 or 8, or the vector of claim 11.

30

14. A method for producing the protein or peptide of any one of claims 6, 7, and 9, comprising culturing the transformant of claim 13 and recovering the expression product.

15. An oligonucleotide comprising the nucleotide sequence of claim 8 (a) or the nucleotide sequence complementary to the complementary strand thereof, wherein said oligonucleotide comprises 15 nucleotides or more.

35

16. Use of the oligonucleotide of claim 15 as a primer for synthesizing a polynucleotide.

17. Use of the oligonucleotide of claim 15 as a probe for detecting a gene.

40

18. An antisense polynucleotide against the polynucleotide of claim 8, or the portion thereof.

19. A method for synthesizing a polynucleotide, the method comprising:

45

a) synthesizing a complementary strand using a cDNA library as a template, and using the primer set of claim 2 or 3, or the primer of claim 16; and

b) recovering the synthesized product.

20. The method of claim 19, wherein the cDNA library is obtainable by oligo-capping method.

50

21. The method of claim 19, wherein the complementary strand is obtainable by PCR.

22. A method for detecting the polynucleotide of claim 8, the method comprising:

55

a) incubating a target polynucleotide with the oligonucleotide of claim 15 under the conditions where hybridization occurs, and

b) detecting the hybridization of the target polynucleotide with the oligonucleotide of claim 15.

23. A database of polynucleotides and/or proteins, the database comprising information on at least one sequence selected from the nucleotide sequences of claim 8 (a) and/or the amino acid sequences of claim 8 (b), or a medium

on which the database is stored.

5

10

15

20

25

30

35

40

45

50

55

Figure 1

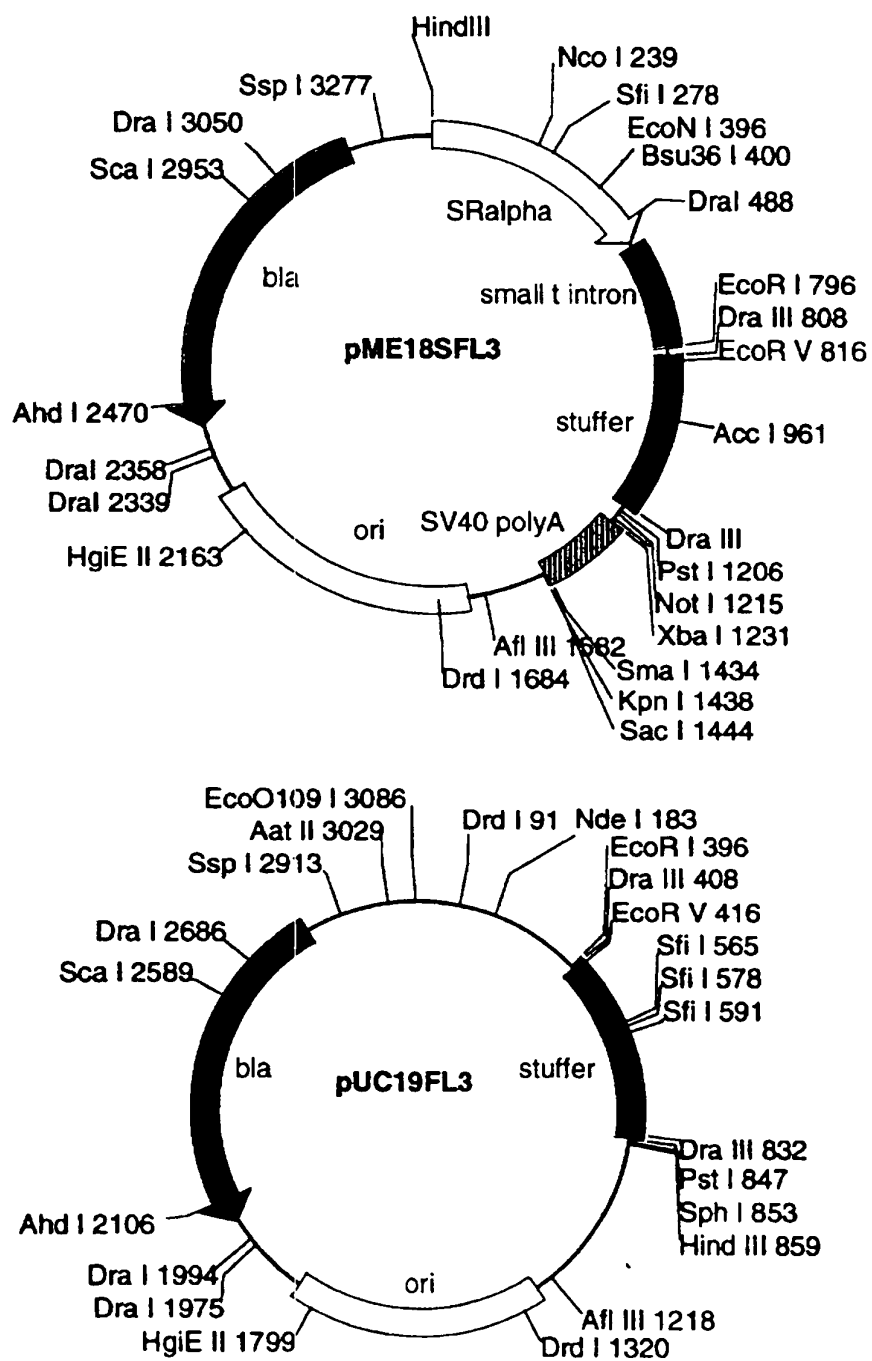


Figure 2

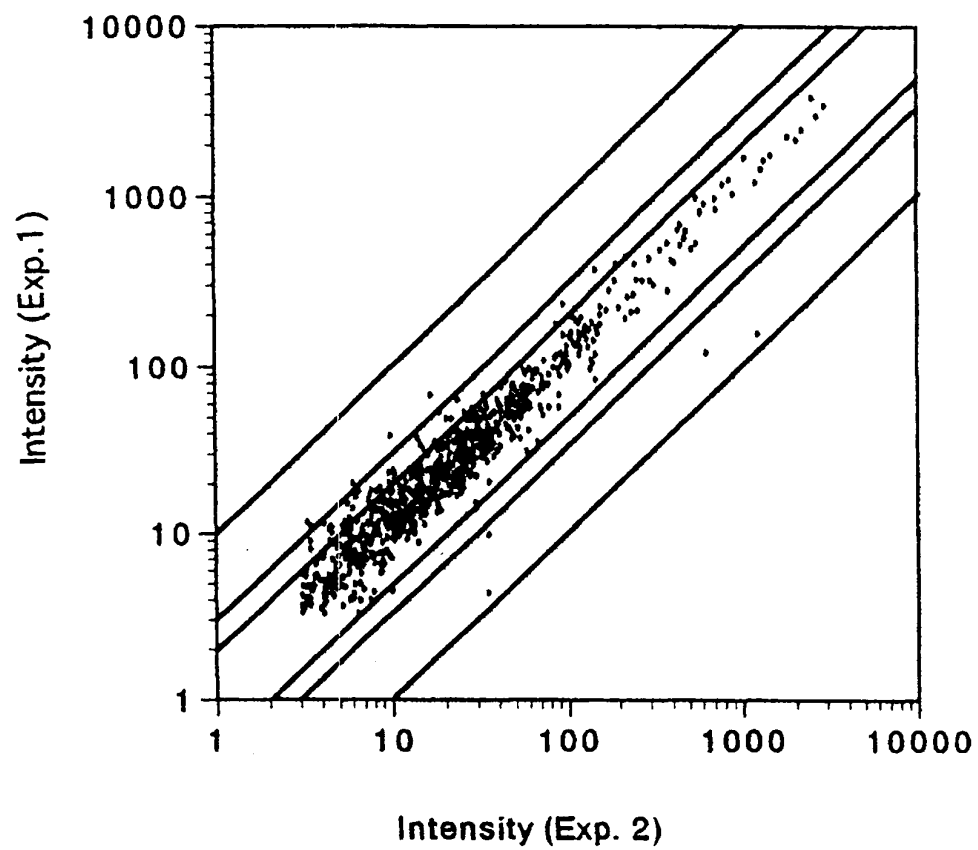
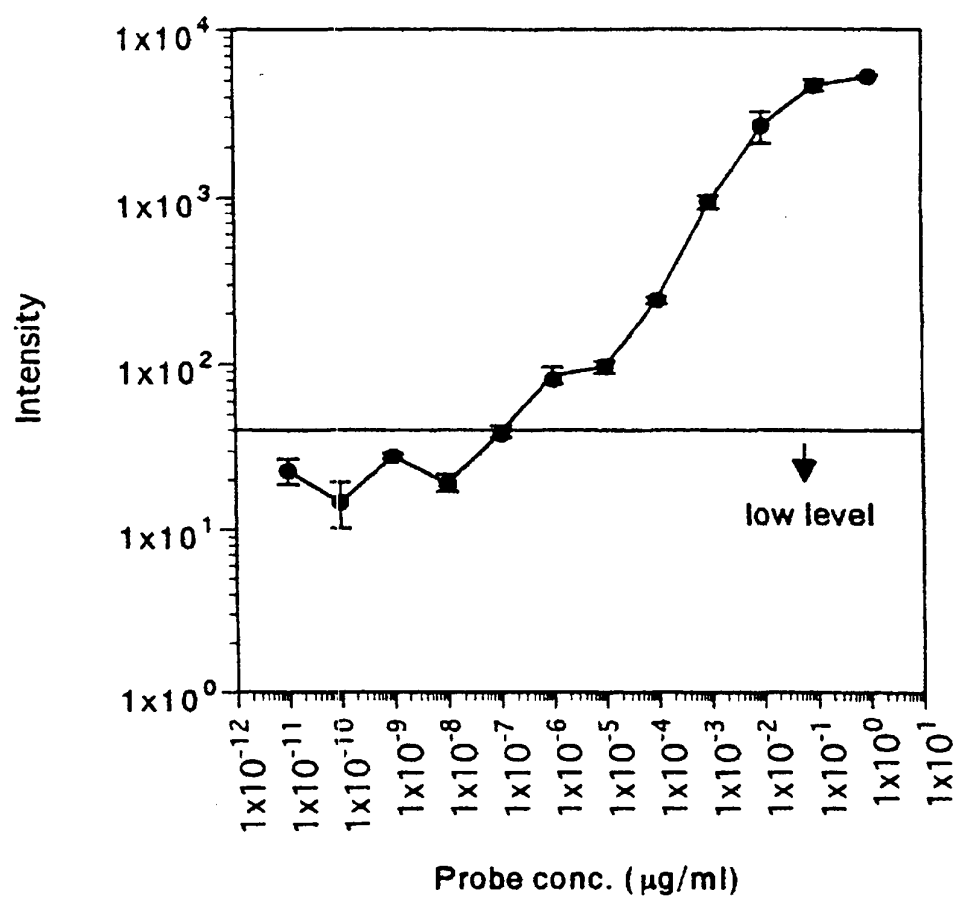


Figure 3



Identifier: AAK94459 cDNA Sequence 2535 BP
 Release Info: Derwent Geneseq Database Release No. 200124; Date released 26-NOV-01
 Database XReference: WPI; 2001-524255/58.;P-PSDB; AAM93529.
 Accession Number: AAK94459
 Patent Title: 830 Primers useful for synthesizing full length cDNA clones and their use in genetic manipulation -
 Patented by: (HELI-) HELIX RES INST.
 Inventor: Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H
 Description: Human full-length cDNA, SEQ ID NO: 3266.
 Patent Number: EP1130094-A2
 Patent Publication Date: 05-SEP-2001
 Modification Date: 06-NOV-2001 (first entry)
 Local Filing: 07-JUL-2000; 2000EP-0114089
 Priority: 08-JUL-1999
 Abstract: The invention relates to primers for synthesising full length cDNA clones. 830 cDNA molecules encoding a human protein have been isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have been determined. Primers for synthesising the full length cDNA are useful for clarifying the function of the protein encoded by the cDNA. The full length clones were obtained by construction of full length enriched cDNA libraries that were synthesised by the oligo-capping method. The primers enable the production of the full length cDNA easily without any special methods. The present sequence is a full length human cDNA of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in CD-ROM format directly from EPO.
 KeyWords: Human;full length cDNA;cDNA synthesis;oligo-capping;ss.
 Organism: Homo sapiens.
 Sequence Composition: Sequence 2535 BP; 482 A; 807 C; 871 G; 375 T; 0 other;
 Sequence: >AAK94459 EP1130094-A2 PA (HELI-) HELIX PR 08-JUL-1999 PF 07-JUL-2000 Human full-length cDNA, SEQ ID NO: 3266. [Homo sapiens.]
 GAGAGCCTGAGCAGGCAGGGCACCTCCAGGTGGGCAGGAGCTACCACTCAGCACCATGAG
 CACCGCCACAGGCCCAGAAGCTGCCCCAAAGCCAAGCGCCAAGTCTATCTATGAGCAGAG
 GAAGCGTTACTCCACAGTTGTTACGGCTGATGTATCCCAGTACCCAGTCAATCACCTGGT
 GACGTTCTGCTGGGTGAGGACGATGGCGTGCATACCGTGGAGGATGCCTCCAGGAAGTT
 GGCCGTCATGGATAGCCAGGGCCGAGTCTGGGCACAGGAGATGCTGCTGCGAGTGTCTCC
 CGACCATGTACGCTGCTCGACCCGGCCTCCAAGGAGGAGCTGGAGTCGTACCCACTGGG
 CGCCATCGTGCGCTGTGACGCGGTGATGCCACCCGGCAGGAGCCGCTCGTTGCTGCTGCT
 CGTGTGCCAGGAACCCGAGCGCGCGCAGCCCGACGTGCACTTCTTCCAGGGCCTGCGCCT
 CGGGGCGGAGCTGATCCGAGAGGACATCCAGGGGGCTCTGCACAATTACCGCTCGGGCCG
 CGGGGAGCGCAGGGCGGGCGGCTCAGGGCCACGCAGGAGGAGTTGCAGCGCGACCGCTC
 GCCC GCCGCTGAGACCCCGCCCTGCAGCGCCGCCGCTCAGTCCGCGCAGTGATCAGCAC
 CGTAGAGCGGGGCGGGGCCGCGGACACCCAGGCGAAGCCCATTCCTCCAGGCAGAGGA
 GGCGCAGAGGCTGAGCCGGTGGGACCTCGAGCAACGCTGACTCGGCTCCCCGGACCT
 GGGTCCCCGGGGTCTGACCTGGCGGTTCTGCGAGGCGGAGCGGGAAGTGGACATCTGAA
 CCACGTGTTGACGACGTAGAGAGCTTTGTATCGAGGCTGCAGAAGTCGGCGGAGGCGGC
 CAGGGTGCTGGAGCACCGGGAACGCGGCCGAGGAGCCGCGCGGGGCGGCTGGGGAGGG
 CTTGCTGACGCTGCGGGCCAAGCCGCCCTCGGAGGCCGAGTACACCGACGTGCTGCAGAA
 GATCAAGTACGCCTTACGCTGCTGGCCCCGGTGCAGCGGCAACATCGCCGACCCCTCCTC
 TCCGGAGCTGTTGCACTTCTTTTCGGGCCTCTGCAGATGATTGTGAACACGTGCGGGGG
 GCCGGAGTTGCGGAGCAGTGTGCGGCGGCCGCATCTGACATCGGATGCCGTGGCGCTGCT
 GCGGGACAACGTCACTCCACGTGA AAACGAGCTCTGGACCTCGCTGGGGGACTCGTGGAC

CCGCCCCGGGCTGGAGCTGTCCCCGGAGGAGGGACCCCCATACAGACCCGAGTTCTTCAG
CGGCTGGGAGCCGCGGTCACTGACCCGCAGAGCCGCGCCTGGGAGGACCCAGTTGAGAA
ACAGCTACAGCACGAGCGGAGGCGCCGGCAGCAAAGCGCCCCCAGGTCGCTGTCAATGG
TCACCGAGACTTGGAGCCAGAATCTGAGCCTCAGCTGGAGTCAGAGACAGCAGGAAAATG
GGTCCTGTGTAATTATGACTTCCAGGCCCCGCAACAGCAGTGAGCTGTCGGTCAAGCAGCG
GGACGTACTGGAGGTCCTGGATGACAGTCGTAAGTGGTGGAAAGGTTCTGGGACCCAGCGGG
GCAGGAGGGATATGTGCCCTACAACATCCTGACACCCTACCCCGGACCCCGGCTGCACCA
CAGCCAAAGCCCTGCCCCGAGCCTGAACAGCACTCCTCCTCCACCACCAGCCCCAGCCCC
GGCCCCACCTCCAGCTCTGGCTCGGCCCCGCTGGGACAGGCCCCGCTGGGACAGCTGCGA
TAGCCTCAACGGCTTGGACCCCGAGGAGAAAGGAGAAATTCTCCAGATGCTCATCGTCAA
CGAGGAACTGCAGGCGCGCCTGGCCCAGGGCCGCTCGGGACCGAGCCGCGCAGTCCCAGG
GCCCCGCGCCCCGGAACCGCAGCTCAGCCCGGGCTCGGACGCCTCCGAGGTCCGCGCCTG
GCTGCAGGCCAAGGGCTTTAGCTCCGGGACCGTGACGCGCTGGGTGTGCTGACCGGGGC
GCAGCTTTTCTCGCTGCAGAGGGAGGAGCTGCGGGCGGTGAGCCCCGAGGAGGGGGCAG
TGTGTACAGCCAGGTCACCGTGACGCTCGCTGCTGGGGGACAAAGAGAAAGTGTGAGA
GCTGGAGGCAGTGATGGAGAAGCAAAAGAAGAAGGTGGAAGGCGAGGTGGAAATGGAGGT
CATTGACCTGCCAGGCGCCCTTCGCAAAGAGTGACGAGGCCCCGTGGGAGAACGGACTC
CTCAGACTCTCCCAATAGCGGAAGTCGATCTTCTGAAGGATGGCCAATCTGCTCCGGCC
CTGGTCTTCCCCCATCCCGGTGGACAGACTTAACGATCCTTGCTGCAGTCCCTCCGGAGA
GGATCTGGACTGGCTGGGAGTGGGGAGGGCGTGGAGACAGTCTACGGAAAGCGCTAGCAG
ACCCCGAGAGGGTGCAGTGGAGCCCTGAGCATTGTAATATGCGGCCCAGCCTATAAACA
GCCTCCGTGCTTAGC